序

众所周知，微积分学和线性代数理论是数学中的两大支柱. 线性代数的重要性在于它考虑了一类简单的数学模型，而大量的理论及应用问题，可以通过“线性化”变成线性代数的问题. 作为基础训练，熟练掌握线性代数的理论和技巧是很有必要的. 这也是我们编写本书的出发点.

本书题为《线性代数与矩阵论》，可以这样去理解：由于引进了直角坐标系，平面上的点便可以用实数对来表示. 反之，实数对的几何意义为：它表示了平面上的点. 所谓平面解析几何学，就是将平面的一些几何问题化为代数问题，确切地说，化为多项式的问题. 和它完全类似，在线性空间中取定一组基后，线性映射可以用矩阵来表示(同样二次型可以用对称方阵来表示). 反之矩阵的几何意义为：它表示了线性空间上的线性映射. 所谓线性代数学，就是或者直接研究线性空间的几何问题，或者将线性空间的一些几何问题化为矩阵问题. 所以线性空间理论和矩阵论实际上是相伴而生的.

本书的特点是偏重矩阵技巧. 这是因为矩阵方法表达具体和明显，并且具有一整套的标准计算技巧和处理办法. 技巧性很强，具有实用价值. 在数值分析、数理统计和经济数学等方面都有重要的应用. 由于矩阵方法着重在技巧，所以初学者不能很快地适应和掌握它. 这只有通过反复计算，才能逐渐熟练. 要深入地理解，就应该概念清楚、运算熟练和灵活运用. 这三个方面是缺一不可的. 正因为如此，书中例题着眼于介绍各种矩阵技巧，并且在每节

后面都附上了习题，其中包括了历届主要单位的研究生考题中有一定难度的题目. 书后附有提示. 但是由于解题方法并不唯一，所以提示仅供参改. 由于题解书不论对学生或老师，都只会促使人们不去思考，丧失通过做题来掌握和熟练数学理论和技巧的功能. 题解书实际上毁了一本习题集，所以我们不希望看到出现本书的题解书.

全书共分十五章. 以一元多项式，多元多项式，行列式，矩阵以及线性方程组求解理论为基础，引进线性空间和线性变换，给出Jordan标准形的几何证明；引进线性函数、双线性函数和多重线性函数，介绍了张量积和外积；讨论了 Euclid 空间和酉空间，二次型分类. 作为应用，讨论了线性不等式和广义逆矩阵；接下去，给出了 Jordan 标准形的代数理论，即 λ 矩阵理论. 在此基础上，引进了方阵函数和复相似下的标准形，处理了矩阵偶在相抵下的标准形理论. 最后，介绍了非负矩阵的基本性质. 书后有两个附录，一个是关于数学归纳法的论述；另一个是给出等价关系的定义和讨论，同时自然地引进了商空间的概念. 贯穿这本书，讲述了各种类型的矩阵标准形理论，它们是初等矩阵论的基本内容.

用本书作教材，建议删去§8.2, §9.3, §9.4, §10.3, §11.5, §13.3, §13.4 以及第十四章和第十五章. 在其他章节中也可适当删去一些内容. 另一方面，要配上若干从教学角度考虑的例题，以及基础训练题. 这是作为教学必须安排的台阶题.

本书适合于作为高年级学生为考研究生的复习参考书，以及低年级高材生的参考书；也适合于作为线性代数课的教学参考书. 我们也希望此书能对科研工作者有较大的参考价值.

多年来，中国科学技术大学数学系用此书为基本教材，很多同志提供了不少宝贵意见. 最近以来，南开大学数学系试点班也用此书作教材，所以在改写过程中，得到顾沛同志的很多帮助. 并且顾

沛同志为此书编写了习题的答案和提示．另外，刘绍学教授和石生明教授的宝贵意见和高教出版社数学编辑张小萍同志的辛勤劳动，都对本书的出版起了很大的作用．在此，一并表示衷心的感谢．由于作者水平有限，书中错误和不妥之处自属难免．还望读者多多提出宝贵意见．

许以超

中国科学院数学研究所

1987年10月7日

目　录

第一章　多项式理论

§1.1　多项式的代数运算

为了书写简便起见，我们先引进**求和符号**，简称**和号**，记作 Σ. 记 n 个数 $a_1, a_2, \cdots, a_n$ 之和 $a_1+a_2+\cdots+a_n$ 为

$$a_1+a_2+\cdots+a_n=\sum_{j=1}^{n} a_j.$$

这个和式中的足码 j 只是表示求和是从 1 加到 n，所以也可以用其他足码来代替．例如，易 j 为 k，即

$$\sum_{j=1}^{n} a_j=\sum_{k=1}^{n} a_k$$

等等．另一方面，由于数的加法有交换律和结合律，所以对 $1,2,\cdots,n$ 的任一排列 $i_1 i_2 \cdots i_n$，上面和式还可以记作

$$\sum_{j=1}^{n} a_j=a_1+\cdots+a_n=a_{i_1}+a_{i_2}+\cdots+a_{i_n}=\sum_{j=1}^{n} a_{i_j},$$

等等．由此可见，不管和式的足码用什么符号，重要的是它实际上表示了 n 个数之和.

现在来考虑**多重和号**，先从二重和号入手．将 mn 个数 a_{jk}，$j=1,2,\cdots,m$; $k=1,2,\cdots,n$ 全部加起来，总和记作 S．下面利用不同的求和方法具体将总和 S 表达出来．先将这 mn 个数排成下表.

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2,n-1} & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3,n-1} & a_{3n} \\ a_{41} & a_{42} & a_{43} & a_{44} & \cdots & a_{4,n-1} & a_{4n} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & a_{m4} & \cdots & a_{m,n-1} & a_{mn} \end{matrix}$$

再按行将 n 个数加起来，它们分别为

$$\sum_{k=1}^{n} a_{1k},\ \sum_{k=1}^{n} a_{2k},\ \cdots,\ \sum_{k=1}^{n} a_{mk},$$

再将这 m 个数加起来，所以有

$$S=(a_{11}+a_{12}+\cdots+a_{1n})+\cdots+(a_{m1}+a_{m2}+\cdots+a_{mn})$$
$$=\sum_{k=1}^{n} a_{1k}+\cdots+\sum_{k=1}^{n} a_{mk}=\sum_{j=1}^{m}\left(\sum_{k=1}^{n} a_{jk}\right).$$

为方便起见，我们记

$$S=\sum_{j=1}^{m}\sum_{k=1}^{n} a_{jk}.$$

也可以按列将 m 个数加起来，它们分别为

$$\sum_{j=1}^{m} a_{j1},\ \sum_{j=1}^{m} a_{j2},\ \cdots,\ \sum_{j=1}^{m} a_{jn},$$

再将这 n 个数加起来，所以有

$$S=(a_{11}+a_{21}+\cdots+a_{m1})+\cdots+(a_{1n}+a_{2n}+\cdots+a_{mn})$$
$$=\sum_{j=1}^{m} a_{j1}+\cdots+\sum_{j=1}^{m} a_{jn}=\sum_{k=1}^{n}\left(\sum_{j=1}^{m} a_{jk}\right).$$

因此证明了

$$S=\sum_{j=1}^{m}\sum_{k=1}^{n} a_{jk}=\sum_{k=1}^{n}\sum_{j=1}^{m} a_{jk}.$$

最后，也可以按对角线将数加起来，它们分别为

$$a_{11}, \quad a_{12}+a_{21}, \quad a_{13}+a_{22}+a_{31}, \quad a_{14}+a_{23}+a_{32}+a_{41}, \cdots$$

由相加的规律可以看出上面每个数 a_{ij} 的足码之和 $i+j$ 分别为定值

$$1+1, \quad 1+2=2+1, \quad 1+3=2+2=3+1,$$
$$1+4=2+3=3+2=4+1, \cdots$$

所以，总和

$$S=a_{11}+(a_{12}+a_{21})+(a_{13}+a_{22}+a_{31})$$
$$+(a_{14}+a_{23}+a_{32}+a_{41})+\cdots$$
$$=\sum_{j+k=2} a_{jk}+\sum_{j+k=3} a_{jk}+\sum_{j+k=4} a_{jk}+\sum_{j+k=5} a_{jk}+\cdots$$

其中足码 j, k 有如下限制：

$$1 \leqslant j \leqslant m, \quad 1 \leqslant k \leqslant n.$$

因此总和

$$S=\sum_{l=2}^{m+n} \sum_{j+k=l} a_{jk}.$$

总之，我们证明了

$$\sum_{j=1}^{m} \sum_{k=1}^{n}=\sum_{k=1}^{n} \sum_{j=1}^{m}=\sum_{l=2}^{m+n} \sum_{j+k=l}.$$

特别，当 $m=n$，则有

$$\sum_{j=1}^{n} \sum_{k=1}^{n}=\sum_{k=1}^{n} \sum_{j=1}^{n},$$

于是可以统一地写成

$$\sum_{j, k=1}^{n}.$$

和上面一样，我们可以引进 p 重和

$$S=\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_p=1}^{n_p} a_{i_1 i_2 \cdots i_p}.$$

它按照关系

$$\sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2}\cdots\sum_{i_p=1}^{n_p}a_{i_1i_2\cdots i_p}$$
$$=\sum_{i_2=1}^{n_2}\sum_{i_3=1}^{n_3}\cdots\sum_{i_p=1}^{n_p}a_{1i_2\cdots i_p}$$
$$+\cdots+\sum_{i_2=1}^{n_2}\sum_{i_3=1}^{n_3}\cdots\sum_{i_p=1}^{n_p}a_{n_1i_2\cdots i_p}$$

由归纳定义出来的. 当 $n_1=n_2=\cdots=n_p$ 时，则可以统一地记成

$$\sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_1}\cdots\sum_{i_p=1}^{n_1}=\sum_{i_1,i_2,\cdots,i_p=1}^{n_1}.$$

对 n 个数 $a_1,a_2,\cdots,a_n$ 之乘积 $a_1a_2\cdots a_n$ 也可简单地记作

$$a_1a_2\cdots a_n=\prod_{j=1}^{n}a_j\ .$$

同样可以用归纳法定义**多重乘积**

$$\prod_{i_1=1}^{n_1}\prod_{i_2=1}^{n_2}\cdots\prod_{i_p=1}^{n_p}a_{i_1i_2\cdots i_p}$$
$$=\left(\prod_{i_2=1}^{n_2}\cdots\prod_{i_p=1}^{n_p}a_{1i_2\cdots i_p}\right)$$
$$\cdots\left(\prod_{i_2=1}^{n_2}\cdots\prod_{i_p=1}^{n_p}a_{n_1i_2\cdots i_p}\right)$$

由于数的乘法也有交换律和结合律,所以符号 Π 和符号 Σ 的性质完全相同.

现在开始讨论一元多项式.

定义 记 x 为未知数,$a_0,a_1,\cdots,a_n$ 为常数,$a_0\neq0$. 则代数式

$$f=f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n=\sum_{j=0}^{n}a_jx^{n-j}$$

称为多项式. n 称为**多项式的次数**,记作 $\deg(f(x))$. 约定常数零为**零多项式**. 它的次数约定为$-\infty$,即有 $\deg(0)=-\infty$. 其它情形也称为**非零多项式**,它们的次数为有限数. 多项式中出现的常数 $a_0, a_1, \cdots, a_n$ 称为**系数**,特别,a_0 称为**首项系数**,a_n 称为**常数项**. 当$a_0, a_1, \cdots, a_n$ 都是复数(实数、有理数、整数)时,多项式 $f(x)$称为**复(实、有理、整)系数多项式**.

定义 两个多项式

$$f(x)=\sum_{j=0}^{n} a_j x^{n-j}, \quad g(x)=\sum_{k=0}^{m} b_k x^{m-k}$$

称为**相等的**,如果 $m=n$,且 $b_j=a_j, 0\leqslant j\leqslant n$.

下面在多项式之间引进代数运算:

1. **加法和减法**

给定多项式

$$f(x)=\sum_{j=0}^{n} a_j x^{n-j}, \quad g(x)=\sum_{j=0}^{n} b_j x^{n-j}$$

这里我们不要求首项系数 $a_0\neq 0, b_0\neq 0$. 则 $f(x)$和$g(x)$ 的和为

$$\begin{aligned} f(x)+g(x) &= \sum_{j=0}^{n} a_j x^{n-j}+\sum_{j=0}^{n} b_j x^{n-j} \\ &= \sum_{j=0}^{n} (a_j+b_j) x^{n-j}. \end{aligned}$$

由**多项式加法**的定义,立即有

(1) $\deg(f(x)+g(x))\leqslant\max(\deg(f(x)), \deg(g(x)))$.

(2) 加法结合律

$$(f(x)+g(x))+h(x)=f(x)+(g(x)+h(x)).$$

(3) 加法交换律

$$f(x)+g(x)=g(x)+f(x).$$

(4) 零多项式 0 有

$$f(x)+0=0+f(x)=f(x).$$

（5）对任一多项式 $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n$，定义多项式

$$-f(x)=(-a_0)x^n+(-a_1)x^{n-1}+\cdots(-a_n),$$

于是有

$$f(x)+(-f(x))=(-f(x))+(f(x))=0.$$

因此可以引进减法：多项式 $f(x)$ 和 $g(x)$ 的差为

$$f(x)-g(x)=f(x)+(-g(x)).$$

2. **乘法**

给定多项式

$$f(x)=\sum_{j=0}^{n}a_jx^{n-j},\quad g(x)=\sum_{k=0}^{m}b_kx^{m-k},$$

则 $f(x)$ 和 $g(x)$ 的乘积为(i) $g=0$, $0\cdot f(x)=f(x)\cdot 0=0$; (ii) $f\neq 0$, $g\neq 0$, $a_0b_0\neq 0$,

$$\begin{aligned}f(x)g(x)&=\left(\sum_{j=0}^{n}a_jx^{n-j}\right)\left(\sum_{k=0}^{m}b_kx^{m-k}\right)\\&=\sum_{j=0}^{n}\sum_{k=0}^{m}a_jb_kx^{n+m-(j+k)}\\&=\sum_{l=0}^{m+n}\left(\sum_{j+k=l}a_jb_k\right)x^{n+m-l}\end{aligned}$$

所以 $f(x)g(x)$ 的首项系数为 $a_0b_0\neq 0$. 由多项式乘法的定义，立即有

（1） $\deg(f(x)g(x))=\deg(f(x))+\deg(g(x))$.

（2） 乘法结合律

$$f(x)(g(x)h(x))=(f(x)g(x))h(x).$$

（3） 乘法交换律

$$f(x)g(x)=g(x)f(x).$$

（4） 零次多项式1有

$$f(x)\cdot 1=1\cdot f(x)=f(x).$$

(5) 乘法消去律

设 $f(x)\neq 0$，且

$$f(x)g(x)=f(x)h(x),$$

则有 $g(x)=h(x)$.

(6) $f(x)g(x)=0$ 当且仅当 $f(x)=0$ 或 $g(x)=0$.

(7) 加乘分配律

$$(f(x)+g(x))h(x)=f(x)h(x)+g(x)h(x).$$

由上面性质可知，多项式的加、减、乘和整数的加、减、乘有完全相同的性质. 而且它们都不能随便作除法，但是有：

定理 1.1.1 设 $g(x)$为非零多项式. 对任一多项式 $f(x)$，则唯一存在多项式 $q(x)$和$r(x)$，使得

$$f(x)=q(x)g(x)+r(x),\quad \deg(r(x))<\deg(g(x)).$$

这里 $q(x)$称为**商式**，$r(x)$称为**余式**.

证 先证存在性. 设 $f(x)=0$，则取 $q(x)=r(x)=0$ 就行了；设 $f(x)\neq 0$，$\deg(g(x)>\deg(f(x))$，则取 $q(x)=0$，$r(x)=f(x)$ 就行了；设 $f(x)\neq 0$，$\deg(g(x))\leqslant\deg(f(x))$. 记

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n,\ g(x)=b_0x^m+b_1x^{m-1}+\cdots+b_m,$$

其中 $a_0b_0\neq 0$，又 $n\geqslant m$. 则多项式

$$f_1(x)=f(x)-b_0{}^{-1}a_0x^{n-m}g(x)$$

的次数小于或等于 $n-1$. 由归纳法假设，便证明了存在多项式 $q_1(x)$，$r(x)$，使得

$$f_1(x)=q_1(x)g(x)+r(x),\quad \deg(r(x))<\deg(g(x)).$$

取 $q(x)=b_0{}^{-1}a_0x^{n-m}+q_1(x)$便证明了分解的存在性.

下面证唯一性. 今若有

$$f(x)=q(x)g(x)+r(x)=q_0(x)g(x)+r_0(x),$$

其中 $\deg\ (r(x))<\deg\ (g(x))$，$\deg(r_0(x))<\deg\ (g(x))$. 所以

$\deg(g(x))>\deg(r(x)-r_0(x))$，且

$$r(x)-r_0(x)=(q_0(x)-q(x))g(x).$$

这证明了当 $r(x)\neq r_0(x)$时，有$q_0(x)\neq q(x)$，且

$$\begin{aligned}\deg g(x)&>\deg(r(x)-r_0(x))\\&=\deg(q_0(x)-q(x))+\deg g(x)\\&>\deg g(x).\end{aligned}$$

所以导出矛盾．因此 $r_0(x)=r(x),q_0(x)=q(x)$．证完．

3．可除性

定义 设 $f(x),g(x)$ 为多项式，其中 $g(x)\neq 0$，如果存在多项式 $q(x)$，使得

$$f(x)=q(x)g(x)$$

则 $g(x)$称为 $f(x)$ 的**因式**，$f(x)$称为 $g(x)$的**倍式**．又称 $g(x)$**除得尽** $f(x)$，记作 $g(x)|f(x)$．当 $g(x)$ **除不尽** $f(x)$ 时，则记作 $g(x)\nmid f(x)$．

引理 1.1.1 除得尽关系有性质：

(1) 传递性

设 $g\neq 0,h\neq 0$，且 $g|f$，$h|g$，则有 $h|f$．

(2) 设 $g\neq 0$，且 $g|f_1$，$g|f_2$，则对任意多项式 h_1，h_2，有 $g|(h_1f_1+h_2f_2)$．

(3) 设 $f\neq 0,g\neq 0$，且 $g|f$，$f|g$，则存在非零常数 c，使得 $g=cf$．

(4) 设 $f\neq 0,\lambda,\mu$ 为非零常数，则 f 有因式 $\lambda,\mu f$．

证 由除得尽的定义可证(1)，(2)，(4)，为了证(3)，还需要多项式乘积的次数公式．证完．

定义 给定多项式 $f_1(x)$和 $f_2(x)$，如果非零多项式 $g(x)$是 $f_1(x)$及 $f_2(x)$的因式，则 $g(x)$称为 $f_1(x)$和 $f_2(x)$的**公因式**．

$f_1(x)$和 $f_2(x)$的公因式$d(x)$称为**最高公因式**，如果 f_1 和 f_2的

任一公因式都是 $d(x)$ 的因式. 当 $f_1(x)$ 和 $f_2(x)$ 不全为零时,首项系数为 1 的最高公因式记作 (f_1, f_2).

引理 1.1.2 给定不全为零的多项式 $f_1(x)$ 和 $f_2(x)$,如果首项系数为 1 的最高公因式存在,则必唯一,且有

$$(f_1, f_2) = (f_2, f_1)$$

证 由最高公因式的定义以及引理 1.1.1 之 (3) 立即可知. 证完.

定理 1.1.2 给定不全为零的多项帐 $f_1(x)$ 和 $f_2(x)$, 则唯一存在首项系数为 1 的最高公因式 (f_1, f_2)

证 由引理 1.1.2,下面只需证明存在性. 为此,引进辗转相除法如下:为方便起见,无妨设 $f_2(x) \neq 0$. 于是有

$$\begin{aligned}
&f_1(x) = q_1(x) f_2(x) + r_1(x),\\
&f_2(x) = q_2(x) r_1(x) + r_2(x),\\
&r_1(x) = q_3(x) r_2(x) + r_3(x),\\
&\cdots\cdots\cdots\\
&r_t(x) = q_{t+2}(x) r_{t+1}(x) + r_{t+2}(x),\\
&\cdots\cdots\cdots
\end{aligned}$$

其中

$$\deg(f_2(x)) > \deg(r_1(x)) > \deg(r_2(x)) > \cdots > \deg(r_t(x)) > \cdots$$

但是 $\deg(f_2(x))$ 为有限数, 所以存在自然数 s, 使得下面算式成立:

$$\begin{aligned}
&f_1(x) = q_1(x) f_2(x) + r_1(x),\\
&f_2(x) = q_2(x) r_1(x) + r_2(x),\\
&r_1(x) = q_3(x) r_2(x) + r_3(x),\\
&\cdots\cdots\cdots\\
&r_{s-2}(x) = q_s(x) r_{s-1}(x) + r_s(x)\\
&r_{s-1}(x) = q_{s+1}(x) r_s(x),
\end{aligned}$$

其中

$$\deg(f_2(x))>\deg(r_1(x))>\deg(r_2(x))>\cdots>\deg(r_s(x))\geqslant 0.$$

记 $r_s(x)$ 的首项系数为 $a\neq 0$，则有

$$(f_1,f_2)=a^{-1}r_s(x).$$

事实上，从上面辗转相除的算式可以看出，对 f_1,f_2 的任一公因式 h，由上往下看，便证明了 $h\mid r_s(x)$. 再由下往上看，可知 $r_s(x)\mid r_{s-1}(x),\cdots,r_s(x)\mid r_1(x),r_s(x)\mid f_2(x),r_s(x)\mid f_1(x)$，即 $r_s(x)$ 为 f_1 和 f_2 的公因式. 由最大公因式之定义可知定理成立. 证完.

定义 给定 m 个不全为零的多项式 $f_1(x),\cdots,f_m(x)$. 非零多项式 $d(x)$ 称为**最高公因式**，如果

(1) $d\mid f_1,\cdots,d\mid f_m$;

(2) 若非零多项式 $d_0(x)$ 有 $d_0\mid f_1,\cdots,d_0\mid f_m$，则 $d_0\mid d$. 首项系数为 1 的最高公因式记作 $(f_1,\cdots,f_m)$.

定理 1.1.3 给定 m 个不全为零的多项式 $f_1,\cdots,f_m$，则它们的首项系数为 1 的最高公因式唯一存在，且有

$$(f_1,\cdots,f_m)=(\cdots((f_1,f_2),f_3),\cdots,f_m).$$

证 用归纳法及最高公因式的定义立即可得. 证完.

辗转相除法还可以导出下面重要性质.

定理 1.1.4 给定两个不全为零的多项式 $f_1(x),f_2(x)$，则存在多项式 $u(x)$ 及 $v(x)$，使得

$$u(x)f_1(x)+v(x)f_2(x)=(f_1,f_2).$$

证 由定理 1.1.2 之证明可知 $(f_1,f_2)=a^{-1}r_s(x)$. 在辗转相除公式中，从上往下，依次消去 $r_1(x),r_2(x),\cdots,r_{s-1}(x)$，便证明了定理. 证完.

由定理 1.1.3 和 1.1.4 立即有

定理 1.1.5 给定 m 个不全为零的多项式 $f_1(x)$, $f_2(x)$, $\cdots$, $f_m(x)$，则存在多项式 $u_1(x),u_2(x),\cdots,u_m(x)$，使得

$$u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_m(x)f_m(x)=(f_1,f_2,\cdots,f_m).$$

习题 1.1

1. 试求一个次数最低的多项式，使得它被 $x^4-2x^3-2x^2+10x-7$ 除时余式为 x^2+x+1；被 $x^4-2x^3-3x^2+13x-10$ 除时余式为 $2x^2-3$.

2. 试求适合条件 $(x^2+1)|f(x)$, $(x^3+x^2+1)|(f(x)+1)$ 的次数最低的多项式 $f(x)$.

3. 试求多项式 $u(x)$ 及 $v(x)$，使得

$$x^m u(x)+(1-x)^n v(x)=1.$$

4. 设 f,g,h 为非零多项式，试证：由 $(f,g)=1$, $(f,h)=1$ 可推出 $(f,gh)=1$. 又 $(hf,hg)=a^{-1}h(f,g)$ 对一切非零多项式 f,g,h 成立，其中 a 为 h 的首项系数.

5. 设 f,g 为非零复(实)系数多项式，试证：(f,g) 也为复(实)系数多项式.

6. 给定两个不全为零的多项式 $f(x)$ 和 $g(x)$. 作多项式集合

$$\{u(x)f(x)+v(x)g(x)\,|\,\forall \text{多项式 } u,v\}$$

试证：这个集合中存在次数最低，且首项系数为 1 的多项式，它就是 $f(x)$ 和 $g(x)$ 的最高公因式 (f,g). 所以它也可以作为最高公因式的等价定义.

7. 给定不全为零的多项式 $f_1(x)$, $f_2(x)$, $f_3(x)$，试证：存在六个多项式 $g_1(x)$, $g_2(x)$, $g_3(x)$, $h_1(x)$, $h_2(x)$, $h_3(x)$，使得三阶行列式

$$\begin{vmatrix} f_1 & f_2 & f_3 \\ g_1 & g_2 & g_3 \\ h_1 & h_2 & h_3 \end{vmatrix}=(f_1,f_2,f_3).$$

§1.2 因 式 分 解

定义 非零多项式 $f(x)$ 和 $g(x)$ 称为**互素的**，如果它们的最大公因式 $(f,g)=1$；否则称为**不互素的**.

定理 1.2.1 非零多项式 f 和 g 互素的必要且充分条件为存在多项式 $u(x)$ 和 $v(x)$，使得 $uf+vg=1$.

证 由定理 1.1.4 及互素的定义可知，当 $(f,g)=1$ 时，必存

在多项式 u,v，使得 $uf+vg=1$. 反之，若存在 u,v 使得 $uf+vg=1$，记 $d=(f,g)$，则有 $d|(uf+vg)$，即 $d|1$. 这证明了 $d=1$，即 $(f,g)=1$，所以 f 和 g 互素. 证完.

引理 1.2.1 非零多项式 f 和 g 及多项式 h 若有 $(f,g)=1$，$f|gh$，则有 $f|h$.

证 由定理 1.2.1，存在多项式 u，v 使得 $uf+vg=1$，于是 $(uh)f+v(gh)=h$. 由条件 $f|gh$，便证明了 $f|h$. 证完.

为了 §3.2 的需要，现在给出

定理 1.2.2 非零多项式 $f(x)$ 和 $g(x)$ 不互素当且仅当存在非零多项式 u_0,v_0，使得

$$u_0f=v_0g,$$

其中

$$0\leqslant\deg(u_0)<\deg(g),\quad 0\leqslant\deg(v_0)<\deg(f).$$

证 设 f 和 g 不互素，于是 $(f,g)=d\neq1$. 记 $f=dv_0,g=du_0$，于是 $u_0f=du_0v_0=v_0g$. 今 $\deg(d)>0$，所以 $0\leqslant\deg(v_0)<\deg(f)$，$0\leqslant\deg(u_0)<\deg(g)$. 反之，若存在 u_0,v_0 有 $u_0f=v_0g$，$0\leqslant\deg(u_0)<\deg(g)$，$0\leqslant\deg(v_0)<\deg(f)$. 下面证 f 和 g 不互素. 用反证法. 若 f 和 g 互素，即 $(f,g)=1$. 由 $f|v_0g$ 及引理 1.2.1，所以 $f|v_0$，这和 $0\leqslant\deg(v_0)<\deg(f)$ 矛盾. 证完.

定义 非零复(实、有理)系数多项式 $p(x)$ 称为**不可约的**，如果除了因式 $\lambda,\mu p(x)$ 外，无其他复(实、有理)系数因式，其中 λ,μ 为非零复(实、有理)数. 非零整系数多项式 $p(x)$ 称为**不可约的**，如果除了 $\pm1,\pm p(x)$ 外，无其他整系数因式. 不是不可约的多项式称为**可约多项式**.

引理 1.2.2 设 $p(x)$ 为不可约多项式，则对任一多项式 $f(x)$，或者 $p|f$，或者 $(p,f)=1$.

证 今若 $(p,f)=1$，则不必证了. 若 $(p,f)=d\neq1$，于是

$\deg(d)>0$，且 $d|p$. 但是 p 不可约，所以 $d=\mu p$，其中 μ^{-1} 为 p 的首项系数. 由 $d|f$ 便证明了 $p|f$. 证完.

引理 1.2.3 设 $p(x)$ 为不可约多项式，则对任两多项式 $f(x)$ 和 $g(x)$，只要 $p|fg$，便有 $p|f$ 或者 $p|g$.

证 今若 $p|f$，则不必证了. 若 $p\nmid f$，由引理 1. 2. 2，所以 $(p,f)=1$. 由引理 1. 2. 1，所以 $p|g$. 证完.

由归纳法立即有

引理 1.2.4 设 $p(x)$ 为不可约多项式，$f_1,f_2,\cdots,f_m$ 为多项式. 设 $p|f_1f_2\cdots f_m$，则 p 必除尽 $f_1,f_2,\cdots,f_m$ 中某个多项式.

定理 1.2.3（唯一析因定理） 设 $f(x)$ 为次数大于零的多项式，则 $f(x)$ 有下列因式分解

$$f(x)=p_1(x)p_2(x)\cdots p_s(x),$$

其中 $p_1,\cdots,p_s$ 为次数大于零的不可约多项式. 且若另有因式分解

$$f(x)=q_1(x)q_2(x)\cdots q_t(x)$$

其中 $q_1,\cdots,q_t$ 为次数大于零的不可约多项式，则 $t=s$，且存在 $1,2,\cdots,s$ 的排列 $i_1i_2\cdots i_s$ 以及 s 个非零常数 $c_1,c_2,\cdots,c_s$，使得 $c_1c_2\cdots c_s=1$，且

$$q_j(x)=c_jp_{i_j}(x),\quad 1\leqslant j\leqslant s.$$

又 $f(x)$ 的次数大于零的因式必为

$$\lambda p_{j_1}(x)p_{j_2}(x)\cdots p_{j_l}(x),$$

其中 λ 为非零常数，$1\leqslant j_1<j_2<\cdots<j_l\leqslant s$，所以 $1\leqslant l\leqslant s$.

证 先证存在性. 对次数作归纳法. 设 $f(x)$ 不可约则定理成立. 设 $f(x)$ 可约. 由定义，$f(x)$ 有因式 $f_1(x)$，$f_2(x)$ 使得 $f(x)=f_1(x)f_2(x)$，且 $0<\deg(f_i)<\deg(f)$，$i=1,2$. 由归纳法，便证明了因式分解的存在性.

再证唯一性，今若 $p_1,\cdots,p_s,q_1,\cdots,q_t$ 为不可约多项式，且有

$$p_1p_2\cdots p_s=q_1q_2\cdots q_t.$$

由 $q_1|p_1p_2\cdots p_s$ 及引理 1.2.4,所以存在 p_{i_1} 使得 $q_1|p_{i_1}$. 但是 p_{i_1} 和 q_1 都是次数大于零的不可约多项式，这证明了存在非零常数 c_1，使得 $q_1=c_1p_{i_1}$. 代回原式,由消去律便得到

$$p_1p_2\cdots p_{i_1-1}p_{i_1+1}\cdots p_s=c_1q_2q_3\cdots q_t.$$

这样依次讨论下去,便证明了 $t\geqslant s$，且存在 1, 2, …, s 的排列 $i_1i_2\cdots i_s$,以及非零常数 $c_1,c_2,\cdots,c_s$,使得

$$1=c_1c_2\cdots c_sq_{s+1}q_{s+2}\cdots q_t.$$

由于 $\deg(q_j)>0, s+1\leqslant j\leqslant t$,这证明了 $t=s$，且 $c_1c_2\cdots c_s=1$. 于是证明了分解的唯一性.

最后，设 $g(x)$ 为次数大于零的多项式，且有 $g(x)|f(x)$. 将 $g(x)$ 分解因式为

$$g(x)=r_1(x)r_2(x)\cdots r_l(x),$$

其中 $r_1(x), r_2(x),\cdots,r_l(x)$ 为次数大于零的不可约多项式. 由 $r_1|g, g|f$,所以 $r_1|f$,即 $r_1|p_1p_2\cdots p_s$. 和唯一性证明相同，有 $r_1=c_1p_{j_1}$. 再依次讨论下去,最后证明了 $g(x)=\lambda p_{j_1}p_{j_2}\cdots p_{j_l}$. 证完.

定义 n 次多项式

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$$

的**导数**定义为多项式

$$f'(x)=na_0x^{n-1}+(n-1)a_1x^{n-2}+\cdots+a_{n-1}.$$

引理 1.2.5 导数有性质

(1) $(f+g)'=f'+g'$;

(2) $(fg)'=f'g+fg'$.

定义 设 $f(x), g(x)$ 为次数大于零的多项式, $e\geqslant 2$ 为自然数. 如果 $g(x)^e|f(x)$，$g(x)^{e+1}\nmid f(x)$，则 $g(x)$ 称为 $f(x)$ 的 e **重因式**. 简称为 $f(x)$ 有**重因式** $g(x)$.

定理 1.2.4 次数大于零的多项式 $f(x)$ 没有重因式的必要且

充分条件为$(f,f')=1$.

证 由定理 1.2.3 可知

$$f(x)=a_0p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_s(x)^{e_s}$$

其中 a_0 为 $f(x)$的首项系数,$p_1(x),p_2(x),\cdots,p_s(x)$为首项系数等于1的不同的不可约多项式,$e_1,e_2,\cdots,e_s$ 为自然数. 由引理 1.2.5,

$$f'(x)=a_0h(x)p_1(x)^{e_1-1}p_2(x)^{e_2-1}\cdots p_s(x)^{e_s-1},$$

其中

$$h(x)=e_1p_1'(x)p_2(x)\cdots p_s(x)+\cdots$$
$$+e_sp_1(x)p_2(x)\cdots p_{s-1}(x)p_s'(x),$$

且由 $\deg(p_j'(x))=\deg(p_j(x))-1, 1\leqslant j\leqslant s$ 不难证明

$$(f(x),f'(x))=p_1(x)^{e_1-1}p_2(x)^{e_2-1}\cdots p_s(x)^{e_s-1}.$$

由 $f(x)$的分解式可知,$f(x)$有重因式当且仅当存在 $e_j>1$,所以 $f(x)$无重因式当且仅当 $e_1=e_2=\cdots=e_s=1$,即$(f,f')=1$. 证完.

定理 1.2.5 设 $f(x)$为次数大于零的多项式,则

$$g(x)=(f(x),f'(x))^{-1}f(x)$$

无重因式,但是 $g(x)$和 $f(x)$有完全相同的不可约因式组.

证 由定理 1.2.4 的证明可知 $f(x)=a_0p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_s(x)^{e_s}$,$g(x)=(f(x),f'(x))^{-1}f(x)=a_0p_1(x)p_2(x)\cdots p_s(x)$. 由定义可知定理成立. 证完.

和最高公因式相伴的,有最低公倍式的概念.

定义 设多项式 $m_0(x)$为非零多项式 $f(x)$和 $g(x)$的倍式,则 $m_0(x)$称为 $f(x)$和 $g(x)$的**公倍式**. $f(x)$和 $g(x)$的公倍式 $m(x)$称为**最低公倍式**,如果 $m(x)$为 $f(x)$和 $g(x)$的公倍式,而且对 $f(x)$和 $g(x)$的任一公倍式 $m_0(x)$,必有 $m(x)\mid m_0(x)$. $f(x)$和 $g(x)$的首项系数为1的最低公倍式记作$[f,g]$.

定理 1.2.6 给定非零多项式 $f(x)$和 $g(x)$,则唯一存在最低

公倍式$[f,g]$，它是

$$[f,g]=\frac{fg}{a_0b_0(f,g)},$$

其中 a_0,b_0 分别为 f,g 的首项系数.

证 为方便起见，无妨设 f 和 g 的首项系数都等于1. 记 $d(x)=(f,g)$，于是

$$f=df_1,\quad g=dg_1,$$

其中$(f_1,g_1)=1$. 于是记

$$m(x)=\frac{fg}{(f,g)}=df_1g_1=f_1g=fg_1.$$

则有 $f|m,g|m$，所以 $m(x)$为 $f(x)$和 $g(x)$的公倍式，且首项系数为1. 另一方面，设 $m_0(x)$为 $f(x)$和 $g(x)$的公倍式，则 $m_0(x)=f(x)l_1(x)=g(x)l_2(x)$，于是 $m_0(x)=d(x)f_1(x)l_1(x)=d(x)g_1(x)l_2(x)$. 这证明了 $f_1(x)l_1(x)=g_1(x)l_2(x)$. 由$(f_1,g_1)=1,g_1|f_1l_1$，用引理 1.2.1 便证明了 $g_1|l_1$，所以 $l_1=g_1l_0$. 代回去，便有

$$m_0(x)=d(x)f_1(x)l_1(x)=d(x)f_1(x)g_1(x)l_0(x)=m(x)l_0(x).$$

这证明了 $m(x)|m_0(x)$. 由定义可知 $m(x)$为 $f(x)$和 $g(x)$的最低公倍式，且首项系数为1，即 $m(x)=[f,g]$. 证完.

定义 给定 r 个非零多项式 $f_1(x)$，$f_2(x)$，$\cdots,f_r(x)$，如果 $m(x)$为 $f_1,f_2,\cdots,f_r$ 的公倍式，且 $f_1,f_2,\cdots,f_r$ 的任一公倍式必为 $m(x)$的倍式，则 $m(x)$称为 $f_1,f_2,\cdots,f_r$ 的**最低公倍式**. 首项系数为1的最低公倍式记作$[f_1,f_2,\cdots,f_r]$.

计算 r 个多项式的最低公倍式用下面定理

定理 1.2.7 对 r 个非零多项式 $f_1(x),f_2(x),\cdots,f_r(x)$，则有 $[f_1,f_2,\cdots,f_r]=[\cdots[[f_1,f_2],f_3],\cdots,f_r]$.

证 用归纳法及最低公倍式的定义立即可证. 证完.

习题 1.2

1. 试求多项式 $f(x)$ 和 $g(x)$，使得 (i) $\deg(f)\neq\deg(g)$；(ii) $f(x)$ 和 $g(x)$ 都是实系数多项式；(iii) $f^2-g^2=x^4+x^3+x^2+x+1$.

2. 给定自然数 k，试证：存在自然数 m，使得

$$(1+x)(1+x^2)\cdots(1+x^{2^k})=1+x+x^2+\cdots+x^m.$$

3. 对什么自然数 n，$1+x+\cdots+x^{n-1}$ 除得尽 $1+x^2+\cdots+x^{2n-2}$.

4. 给定常数 a，定义映射 $f\to D(f)$，其中 f 为多项式，$D(f)$ 为数，且有

(1) $D(\lambda f+\mu g)=\lambda D(f)+\mu D(g)$，$\lambda,\mu$ 为常数，

(2) $D(fg)=f(a)D(g)+g(a)D(f)$，

其中 f,g 为多项式. 试证：存在常数 b，使得 $D(f)=bf'(a)$.

5. 设 $u_i=a_ix+b_i$，$i=1,2,3$ 为实系数多项式，有

$$u_1^n+u_2^n=u_3^n,$$

其中 n 为任意给定的自然数，且 $n\geqslant 2$. 试证：存在多项式 $p(x)=Ax+B$，使得 $u_i=c_ip(x)$，$i=1,2,3$，其中 c_1,c_2,c_3 为常数.

6. 记

$$f(x)=(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)=x^n-\sigma_1x^{n-1}+\cdots+(-1)^n\sigma_n$$

记

$$s_k=\alpha_1^k+\alpha_2^k+\cdots+\alpha_n^k,\quad k=0,1,2,\cdots.$$

试证：

$$g_k(x)=x^{k+1}f'(x)-(s_0x^k+s_1x^{k-1}+\cdots+s_k)f(x),\quad k=0,1,2,\cdots,$$

有 $\deg(g_k)<\deg(f)$. 并用上式推出 Newton 公式

$$s_k-\sigma_1s_{k-1}+\sigma_2s_{k-2}-\cdots+(-1)^{k-1}\sigma_{k-1}s_1+(-1)^k\sigma_ks_0=0,\ 1\leqslant k<n;$$

$$s_k-\sigma_1s_{k-1}+\sigma_2s_{k-2}-\cdots+(-1)^n\sigma_ns_{k-n}=0,\ k\geqslant n$$

7. 设 $f(x),g(x),h(x)$ 为非零多项式，试证：

$$[f,(g,h)]=([f,g],[f,h]).$$

§1.3 整系数多项式

设 $f(x)$ 为有理系数多项式，将所有系数的分母通分，便可写成

$$f(x)=\frac{a}{b}g(x)$$

其中 a,b 为整数，且$(a,b)=1$，即 a/b 为既约分数. 又

$$g(x)=c_0x^n+c_1x^{n-1}+\cdots+c_n$$

为整系数多项式，其中 c_0为正整数. 又$(c_0,c_1,\cdots,c_n)=1$. 由此可知有理系数多项式的因式分解和整系数多项式的因式分解密切相关.

定义 非零整系数多项式$g(x)=\sum_{j=0}^{n}c_jx^{n-j}$称为**本原多项式**，如果$(c_0,c_1,\cdots,c_n)=1$，即 $c_0,c_1,\cdots,c_n$ 互素.

用这个定义，上面讨论可写成

引理 1.3.1 任一非零有理系数多项式为一个有理数和一个本原多项式的乘积.

关于整系数多项式分解为整系数多项式乘积的因式分解理论中，重要的是

引理 1.3.2(Gauss 引理) 任两本原多项式的乘积仍为本原多项式

证 给定两个本原多项式

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n,\quad g(x)=b_0x^m+b_1x^{m-1}+\cdots+b_m,$$

于是

$$(a_0,a_1,\cdots,a_n)=1,(b_0,b_1,\cdots,b_m)=1.$$

而

$$\begin{aligned}f(x)g(x)&=\sum_{j=0}^{n}a_jx^{n-j}\sum_{k=0}^{m}b_kx^{m-k}\\&=\sum_{l=0}^{n+m}\Big(\sum_{j+k=l}a_jb_k\Big)x^{n+m-l}\\&=\sum_{l=0}^{n+m}c_lx^{n+m-l}\end{aligned}$$

问题化为要证明$(c_0,c_1,\cdots,c_{n+m})=1$. 设若不然，则存在素数 p，使

得 $p|c_l, 0\leqslant l\leqslant n+m$. 今 $p|c_0$，即 $p|a_0b_0$. 由引理 1.2.3，$p|a_0$ 或者 $p|b_0$. 无妨设 $p|a_0$，由 $(a_0,a_1,\cdots,a_n)=1$，所以存在指标 r，$0\leqslant r<n$，使得 $p|a_0, p|a_1,\cdots,p|a_r$，但是 $p\nmid a_{r+1}$. 今由

$$c_{r+1}=a_0b_{r+1}+a_1b_r+\cdots+a_rb_1+a_{r+1}b_0$$

及 $p|c_{r+1}$，这证明了 $p|a_{r+1}b_0$，所以 $p|b_0$. 由 $(b_0,b_1,\cdots,b_m)=1$，所以存在指标 s，$0\leqslant s<m$，使得 $p|b_0,\cdots,p|b_s, p\nmid b_{s+1}$. 由

$$c_{r+s+2}=a_0b_{r+s+2}+a_1b_{r+s+1}+\cdots+a_{r+s+1}b_1+a_{r+s+2}b_0$$

及 $p|c_{r+s+2}$，所以 $p|a_{r+1}b_{s+1}$，这和 $p\nmid a_{r+1}, p\nmid b_{s+1}$ 矛盾. 证完.

引理 1.3.3 任一整系数多项式必为一个整数和一个本原多项式的乘积. 所以次数大于零的整系数不可约多项式必为本原多项式. 次数等于零的整系数不可约多项式为 $\pm p$，其中 $p=1$ 或为素数.

证明极易，在此略去.

定理 1.3.1 设整系数本原多项式 $f(x)$ 的次数大于零，则 $f(x)$ 不可约当且仅当它作为有理系数多项式仍不可约.

证 设本原多项式 $f(x)$ 作为有理系数多项式，它不可约，自然它作为整系数多项式也不可约. 反之，若本原多项式 $f(x)$ 作为整系数多项式，它不可约. 但作为有理系数多项式，它可约. 于是 $f(x)=g(x)h(x)$，其中 $g(x)=\frac{a}{b}g_1(x)$，$h(x)=\frac{c}{d}h_1(x)$，而 $\frac{a}{b}$ 及 $\frac{c}{d}$ 为既约分数，$g_1(x)$，$h_1(x)$ 为本原多项式，又 $0<\deg g(x)<\deg f(x)$. 由引理1.3.2，$g_1(x)h_1(x)$ 仍为本原多项式，而 $f(x)=\frac{ac}{bd}(g_1(x)h_1(x))$. 由 $f(x)$ 及 $g_1(x)h_1(x)$ 为本原多项式，便证明了 $bd=\pm ac$. 所以 $f(x)=\pm g_1(x)h_1(x)$. 由 $0<\deg g_1(x)<\deg f(x)$，便证明了 $f(x)$ 作为整系数多项式，它可约. 这和假设矛盾，所以证明了定理.

定理 1.3.2 (Eisenstein 不可约判别法) 给定次数大于零的本原多项式

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n,\quad a_0\neq 0.$$

如果存在素数 p 使得

$$p\nmid a_0,\quad p\mid a_1,\cdots,\quad p\mid a_n,\quad p^2\nmid a_n,$$

则 $f(x)$ 为整系数不可约多项式.

证 用反证法，若 $f(x)$ 为整系数可约多项式，则存在整系数多项式 $g(x)=\sum_{j=0}^{r}b_jx^{r-j}$, $h(x)=\sum_{k=0}^{n-r}c_kx^{n-r-k}$，使得 $0<\deg g(x)<n$，即 $0<r<n$，且 $f(x)=g(x)h(x)$. 于是

$$a_l=\sum_{j+k=l}b_jc_k,\quad 0\leqslant l\leqslant n.$$

今 $a_n=b_rc_{n-r}$, $p\mid a_n$, $p^2\nmid a_n$，于是 $p\mid b_r$ 或者 $p\mid c_{n-r}$. 为方便起见，无妨设 $p\mid b_r$. 由 $p^2\nmid b_rc_{n-r}$，所以 $p\nmid c_{n-r}$. 今若 $p\mid b_r, p\mid b_{r-1}$, $\cdots$, $p\mid b_{r-i+1}$，则由

$$p\mid a_{n-i}=b_rc_{n-r-i}+b_{r-1}c_{n-r-i+1}+\cdots+b_{r-i+1}c_{n-r-1}+b_{r-i}c_{n-r},$$

$i=1, 2,\cdots,r$，便证明了 $p\mid b_{r-i}, i=1,2,\cdots,r$，特别 $p\mid b_0$. 由于 $a_0=b_0c_0$ 及 $p\mid b_0$ 可知 $p\mid a_0$，于是和假设矛盾. 证完.

习题 1.3

1. 试证下列整系数多项式不可约:

(i) $x^{p-1}+x^{p-2}+\cdots+x+1$, p 为素数;

(ii) x^5-x^2+1;

(iii) $(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)-1$，其中 α_1, α_2, $\cdots$, α_n 为 n 个不同整数.

2. 试证: 本原多项式 $x^{10}\pm x^5+1$ 可约.

3. 给定 r 个不同素数 $p_1, p_2, \cdots, p_r$，试证: $\sqrt[m]{p_1p_2\cdots p_r}$ 为无理数，其中 $m\geqslant 2$ 为自然数.

§1.4 多项式的根

定义 给定多项式 $f(x)$，数 α 称为多项式 $f(x)$ 的根，如果 $f(\alpha)=0$.

显然，所有数都是零多项式的根，又零次多项式无根，下面讨论次数大于零的多项式的根.

定理 1.4.1 $f(x)$以 α 为根当且仅当$(x-\alpha)\mid f(x)$.

证 由定理 1.1.1，用 $x-\alpha$ 除 $f(x)$，则有

$$f(x)=(x-\alpha)q(x)+c,$$

其中 c 为常数，$q(x)$为商式. 取 $x=\alpha$，则有 $c=f(\alpha)$. 所以证明了

$$f(x)=(x-\alpha)q(x)+f(\alpha).$$

因此 $f(\alpha)=0$ 当且仅当$(x-\alpha)\mid f(x)$. 证完.

定理 1.4.2 n 次($n\geqslant 1$)多项式至多有 n 个不同的根.

证 设 $f(x)$为 n 次多项式. 设 α_1 为 $f(x)$的根，于是 $f(x)=(x-\alpha_1)f_1(x)$，其中 $f_1(x)$为 $n-1$ 次多项式. 显然 $f_1(x)$的根仍为 $f(x)$的根. 用归纳法，已知一次多项式恰好有一个根. 设 $n-1$次多项式至多有 $n-1$ 个不同根，由 $f(x)=(x-\alpha_1)f_1(x)$可知 n 次多项式 $f(x)$至多有 n 个不同根. 证完.

定理 1.4.3 设存在 $n+1$ 个不同的数 $b_0, b_1, \cdots, b_n$，使得

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n$$

有 $f(b_j)=0, 0\leqslant j\leqslant n$，则有 $a_0=a_1=\cdots=a_n=0$.

证 用反证法，若 $f(x)\neq 0$，则无妨设 $a_0\neq 0$, $\deg f(x)=n$. 于是 $f(x)$有 $n+1$ 个不同根$b_0, b_1, \cdots, b_n$，由定理 1.4.2 便导出矛盾. 证完.

定理 1.4.4 设多项式 $f(x)$ 和 $g(x)$ 的次数都不超过 n，则

$f(x)=g(x)$ 当且仅当存在 $n+1$ 个不同的数 $b_0,b_1,\cdots,b_n$，使得 $f(b_j)=g(b_j), j=0,1,\cdots,n$.

证 考虑多项式 $F(x)=f(x)-g(x)$，由定理 1.4.3 便证明了定理. 证完.

下面分别在不同的数的范围内讨论根的存在性及求根方法.

(1) 复数情形

定理 1.4.5（代数基本定理） 设 $n\geqslant 1$，则 n 次复系数多项式必有复根.

证明可以在复变函数论中学到，在此略去.

定理 1.4.6 设 $n\geqslant 1$，则 n 次复系数多项式必可分解为 n 个一次因式之乘积. 于是 n 次复系数多项式恰有 n 个复根.

证 由定理 1.4.1 以及归纳法立即可证. 证完.

定理 1.4.7 不可约复系数多项式是且只是零次复系数多项式及一次复系数多项式.

证 由不可约多项式的定义及定理 1.4.6 立即可证. 证完.

定义 复系数多项式 $f(x)$ 若有复根 α，且若 $(x-\alpha)^k|f$，$(x-\alpha)^{k+1}\nmid f$，则 α 称为 $f(x)$ 的 k 重根.

定理 1.4.8 复系数多项式 $f(x)$ 和 $g(x)=(f,f')^{-1}f$ 有完全相同的根，但是 $g(x)$ 无重根. 因此复系数多项式 $f(x)$ 无重根当且仅当 $(f,f')=1$.

证 由定理 1.2.5 以及定理 1.4.6，1.4.1 立即可证. 证完.

(2) 实数情形

记
$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n$$
为复系数多项式. 于是 $f(x)$ 为实系数多项式当且仅当 $\bar{a}_j=a_j$，$0\leqslant j\leqslant n$，即

$$\overline{f(x)}=f(\bar{x}).$$

定理 1.4.9 实系数多项式 $f(x)$若有根 α，则必有根 $\bar{\alpha}$. 若 α 不是实数，则 $f(x)$ 有实因式

$$(x-\alpha)(x-\bar{\alpha})=x^2-(\alpha+\bar{\alpha})x+|\alpha|^2.$$

证明由定理 1.4.1 可知.

显然当 α 不是实数，即 $\alpha\neq\bar{\alpha}$ 时，$(x-\alpha)(x-\bar{\alpha})=x^2-(\alpha+\bar{\alpha})x+|\alpha|^2$ 为实系数不可约多项式，这时它的判别式 $(\alpha+\bar{\alpha})^2-4|\alpha|^2=(\alpha-\bar{\alpha})^2<0$. 结合定理 1.4.6 及定理 1.4.9，便证明了

定理 1.4.10 实系数不可约多项式是且只是零次、一次及二次实系数多项式，在二次的情形，还要求判别式小于零.

由于实系数多项式的根和它的共轭根成对出现，所以也证明了

定理 1.4.11 当 $n\geqslant1$ 时，n 次实系数多项式恰有 $2r$ 个复根 $\alpha_1,\cdots,\alpha_r,\bar{\alpha}_1,\cdots,\bar{\alpha}_r$，其中 $\alpha_1,\cdots,\alpha_r$ 的虚部不等于零；且有 $n-2r$ 个实根 $c_1,c_2,\cdots,c_{n-2r}$. 所以它能分解为下面不可约因式之乘积

$$f(x)=a_0[(x-\alpha_1)(x-\bar{\alpha}_1)]\cdots[(x-\alpha_r)(x-\bar{\alpha}_r)]$$
$$\cdot(x-c_1)(x-c_2)\cdots(x-c_{n-2r}).$$

当然，具体去求一个已知实系数多项式的实根是较为困难的事. 已经发现许多方法去求近似根. 这些方法属于计算数学的一个分支. 下面给出一种经典的办法来求实根的近似值，这就是著名的 Sturm定理.

定义 给定实系数非零多项式 $f(x)$，对多项式 $f(x)$及 $f'(x)$ 作辗转相除，记作 $Q_0(x)=f(x)$, $Q_1(x)=f'(x)$,

且

$$Q_0(x)=q_1(x)Q_1(x)-Q_2(x),$$
$$Q_1(x)=q_2(x)Q_2(x)-Q_3(x),$$
$$\cdots\cdots$$
$$Q_{s-2}(x)=q_{s-1}(x)Q_{s-1}(x)-Q_s(x),$$
$$Q_{s-1}(x)=q_s(x)Q_s(x),$$

其中

$$\deg Q_1(x) > \deg Q_2(x) > \cdots > \deg Q_s(x) \geqslant 0,$$

则多项式序列

$$Q_0(x), Q_1(x), Q_2(x), \cdots, Q_s(x)$$

称为**Sturm 序列**.

定义 给定 $s+1$ 个实数 $c_0, c_1, \cdots, c_s$. 从左向右观察，如果 $c_i=0$，则除去；如果 $c_i \neq 0, c_{i+1}=\cdots=c_{j-1}=0, c_j \neq 0$，且 c_i 和c_j 符号相同(即同时为正数或为负数)，则称**无变号**；当 c_i 和 c_j 的符号相反(即其中有一为正数，另一为负数)，则称有**一个变号**. 变号数之总和称为该序列之**变号总数**.

定义 给定实系数多项式 $f(x)$. 记它的 Sturm 序列为$Q_0(x)$, $Q_1(x), \cdots, Q_s(x)$. 对任一实数 c，记序列 $Q_0(c), Q_1(c), \cdots, Q_s(c)$ 的变号总数为 $w(c)$. 于是得到实轴上函数 $w(x)$，称为 Sturm 序列 $Q_0(x), Q_1(x), \cdots, Q_s(x)$的**变号函数**.

引理 1.4.1 设非零多项式 $f(x)$ 为无重根的实系数多项式，即$(f, f')=1$. 则 $f(x)$的 Sturm 序列有性质

(1) $Q_s(x)$为零次多项式；

(2) 相邻两多项式 $Q_j(x)$, $Q_{j+1}(x)$无公共实根；

(3) 如果实数 c 有 $Q_j(c)=0$，则$Q_{j-1}(c)=-Q_{j+1}(c)$，$1 \leqslant j \leqslant s-1$；

(4) 如果实数 c 有 $Q_0(c)=0$，则存在 $\varepsilon>0$，使得当 $c-\varepsilon<x<c$ 时有 $Q_0(x)Q_1(x)<0$；当 $c<x<c+\varepsilon$ 时有 $Q_0(x)Q_1(x)>0$.

证 由于 $Q_s(x)$为 f 及 f' 之最大公因式及 $(f, f')=1$，所以 $Q_s(x)$为常数，即(1)成立. 若 $Q_j(x), Q_{j+1}(x)$ 有公共根 α，则由 Sturm 序列之定义可知 $Q_{j+2}(x), \cdots, Q_s(x)$有公共根 α，这和(1)矛盾，所以(2)成立. 今若 $1 \leqslant j \leqslant s-1$，于是有

$$Q_{j-1}(x)=q_j(x)Q_j(x)-Q_{j+1}(x).$$

如果实数 c 有$Q_j(c)=0$，自然 $Q_{j-1}(c)=-Q_{j+1}(c)$，此即(3)成立.

最后，如果 c 为$Q_0(x)$的根，即$Q_0(c)=0$. 由$(f,f')=1$ 可知$Q_1(c)\neq 0$. 但是 $Q_1(x)$为连续函数，所以存在 $\varepsilon>0$，使得当$|x-c|<\varepsilon$，有 $Q_1(x)=f'(x)\neq 0$. 当 $c-\varepsilon<x<c+\varepsilon$ 时有 $f'(x)>0$，则 $f(x)$ 单调递增，但是 $Q_0(c)=f(c)=0$，所以在 $c-\varepsilon<x<c$ 时有 $f(x)<0$，即 $f(x)f'(x)<0$；在 $c<x<c+\varepsilon$ 时有 $f(x)>0$，即 $f(x)f'(x)>0$. 当 $c-\varepsilon<x<c+\varepsilon$ 时有 $f'(x)<0$，则 $f(x)$单调递减，所以在 $c-\varepsilon<x<c$时有 $f(x)>0$，即$f(x)f'(x)<0$；在$c<x<c+\varepsilon$ 时有 $f(x)<0$，即 $f(x)f'(x)>0$. 而 $f(x)=Q_0(x)$，$f'(x)=Q_1(x)$ 便证明了(4)成立. 证完.

定理 1.4.12(Sturm 定理) 设 $f(x)$ 为无重根的实系数非零多项式，即$(f,f')=1$. 设实数 α,β 不是 $f(x)$的实根，且 $\alpha<\beta$. 则实系数多项式 $f(x)$在区间(α,β)中恰有 $w(\alpha)-w(\beta)$个实根.

证 由于Sturm 序列 $Q_0(x),Q_1(x),\cdots,Q_s(x)$由多项式构成，所以当 x 增大时，如果不经过每个多项式 $Q_j(x)$, $0\leqslant j<s$ 的根时，变号函数 $w(x)$的值不改变. 所以为了证明定理，只要证明当 x 增大时，如果不经过 $Q_0(x)$的根时，变号函数的值也不改变，而经过 $Q_0(x)$的一个根时，变号函数的值减少 1. 我们证明如下：

设实数 c 不是 $Q_0(x)$的根，是某个 $Q_j(x)$的根. 由引理1.4.1 的性质(1)，j 为 1, 2, $\cdots$, $s-1$ 之一. 由性质(3)，$Q_{j-1}(c)=-Q_{j+1}(c)$. 由性质(2)，$Q_{j-1}(c)Q_{j+1}(c)\neq 0$. 所以当$Q_{j-1}(c)>0$，有下表

	$Q_{j-1}(x)$	$Q_j(x)$	$Q_{j+1}(x)$
$c-\varepsilon<x<c$	+	$\pm$	−
$x=c$	+	0	−
$c<x<c+\varepsilon$	+	$\mp$	−

因此当 x 从 $c-\varepsilon$ 到 $c+\varepsilon$ 变化时，$Q_{j-1}(x),Q_j(x),Q_{j+1}(x)$ 的变号

数始终为1；当 $Q_{j-1}(c)<0$，有下表

	$Q_{j-1}(x)$	$Q_j(x)$	$Q_{j+1}(x)$
$c-\varepsilon<x<c$	$-$	$\pm$	$+$
$x=c$	$-$	0	$+$
$c<x<c+\varepsilon$	$-$	$\mp$	$+$

因此当 x 从 $c-\varepsilon$ 增大到 $c+\varepsilon$ 的过程中，$Q_{j-1}(x)$，$Q_j(x)$，$Q_{j+1}(x)$的变号数也始终为1. 所以取 $j=1,2,\cdots,s-1$，便证明了当 x 增大且不经过 $Q_0(x)$的根时，变号函数 $w(x)$的值不改变.

余下讨论当 c 为$Q_0(x)$的根的情形，由性质(2)可知$Q_1(c)\neq0$，所以存在 $\varepsilon>0$，使当$|x-c|<\varepsilon$ 时 $Q_1(x)$ 不变号，由性质(4)可知，有下表

	$Q_0(x)$	$Q_1(x)$
$c-\varepsilon<x<c$	$\pm$	$\mp$
$x=c$	0	$\mp$
$c<x<c+\varepsilon$	$\mp$	$\mp$

因此当 x 从 $c-\varepsilon$ 增大到 $c+\varepsilon$ 的过程中，$Q_0(x)$，$Q_1(x)$的变号数由1变为0，即在 $c-\varepsilon<x<c$ 时有一个变号数，在 $c<x<c+\varepsilon$ 时没有变号数. 总之，证明了当通过 $Q_0(x)=f(x)$的根 c 时，变号函数 $w(x)$的值减少一个. 这证明了定理. 证完.

(3) 有理数的情形

由引理 1.3.1，有理系数多项式的有理数根为相应本原多项式的有理数根，反之亦然. 所以问题化为：如何求本原多项式的有理数根. 这见下面(4).

(4) 整数的情形

给定本原多项式

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n, a_0\neq 0, n\geqslant 1.$$

下面介绍求 $f(x)$的整数根及有理数根的试验法.

设 m 为 $f(x)$的整数根,即有

$$f(m)=a_0m^n+a_1m^{n-1}+\cdots+a_n=0.$$

于是 $m\mid a_n$. 因此可先求出 a_n 的所有因子, 将每个因子代入 $f(x)$ 中,计算它何时为零,从而可求出 $f(x)$的所有整数根.

设$\frac{m}{q}$为既约分数, 即有$(m,q)=1$. 设$\frac{m}{q}$为 $f(x)$ 的有理数根, 即有

$$f\left(\frac{m}{q}\right)=a_0\left(\frac{m}{q}\right)^n+a_1\left(\frac{m}{q}\right)^{n-1}+\cdots+a_n=0,$$

$$a_0m^n+a_1m^{n-1}q+\cdots+a_nq^n=0.$$

于是 $q\mid a_0, m\mid a_n$. 因此可先求出 a_0 及 a_n 的所有因子, 分别记作 q, m,得到整数对$\{q,m\}$. 我们先筛去$(q,m)\neq 1$ 的整数对,余下的整数对可定义一个既约分数$\frac{m}{q}=x_0$,再将每个 x_0 代入 $f(x)$中, 计算它何时为零,从而可求出 $f(x)$的所有有理数根.

习题 1.4

1. 设 $f(x), p(x)$为非零多项式,且有公根 α. 试证: 当 $p(x)$为不可约多项式时有 $p(x)\mid f(x)$.

2. 求证: 下列多项式都以 1 为三重根:

(1) $x^{2n}-nx^{n+1}+nx^{n-1}-1$;

(2) $(n-2m)x^n-nx^{n-m}+nx^m-(n-2m), n>m$;

(3) $x^{2n+1}-(2n+1)x^{n+1}+(2n+1)x^n-1$.

3. 设 x_0 为多项式 $f_1(x)f_2'(x)-f_2(x)f_1'(x)$的 k 重根, 且 $f_1(x_0)f_2(x)\neq$

$f_2(x_0)f_1(x)$. 试证: x_0 为多项式 $f_1(x_0)f_2(x)-f_2(x_0)f_1(x)$ 的 $k+1$ 重根. 反之亦然.

4. 求自然数 n, 使得 $(x+1)^n-x^n-1$ 无重根.

5. 试证:

$$f(x)=1+\frac{1}{1!}x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n$$

无重根.

6. 试证: $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n$ 能被 $(x-1)^{k+1}$ 除尽的必要且充分条件为 $k+1\leqslant n$, 且

$$\sum_{j=0}^{n} j^t a_j=0,\ t=0,1,\cdots,k.$$

7. 试证: $f'|f$ 当且仅当 $f=a_0(x-a)^n$.

8. 设 $F(x)$ 为实系数多项式, 由数学分析可知当 $F'(x_0)=0$, $F''(x_0)<0$ 时 $F(x)$ 在 $x=x_0$ 达极大值; 当 $F'(x_0)=0, F''(x_0)>0$ 时 $F(x)$ 在 $x=x_0$ 达极小值. 用这个结果, 试证明: 设 $f(x)$ 为 n 次实系数多项式, 且对一切实数 x, 有 $f(x)\geqslant 0$, 则对一切实数 x, 有

$$f(x)+f'(x)+f''(x)+\cdots+f^{(n)}(x)\geqslant 0.$$

9. 设多项式 $f(x), g(x), h(x)$ 有

$$f(x^5)+xg(x^5)+x^2h(x^5)=(x^4+x^3+x^2+x+1)k(x).$$

试证: f, g, h 以 $x-1$ 为公因式.

10. 试证 Lagrange 插值定理: 给定 $2n$ 个数 $\alpha_1,\alpha_2,\cdots,\alpha_n,a_1, a_2, \cdots, a_n$, 则唯一存在一个次数不超过 $n-1$ 的多项式 $f(x)$, 它有

$$f(\alpha_i)=a_i, 1\leqslant i\leqslant n.$$

事实上,

$$f(x)=\sum_{j=1}^{n}a_j\frac{(x-\alpha_1)\cdots(x-\alpha_{j-1})(x-\alpha_{j+1})\cdots(x-\alpha_n)}{(\alpha_j-\alpha_1)\cdots(\alpha_j-\alpha_{j-1})(\alpha_j-\alpha_{j+1})\cdots(\alpha_j-\alpha_n)}.$$

11. 设多项式 $f(x)$ 被 $x-1, x-2, x-3$ 除后, 余式分别为 4, 8, 16. 试求 $f(x)$ 被 $(x-1)(x-2)(x-3)$ 除后之余式.

12. 设 f 为实系数多项式. 试证: f 的根都是实根当且仅当 f^2 不能表成两个次数不同的实系数多项式的平方和.

13. 试证: 实系数多项式

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-3}x^3+x^2+x+1,\ n\geqslant 3$$

必有复根，它的虚部不等于零.

14. 设$f(x)$为实系数多项式，则$f(x)$关于一切实数x_0，总有$f(x_0)\geqslant 0$当且仅当存在实系数多项式$g(x)$，$h(x)$，使得$f=g^2+h^2$.

15. 设实系数多项式$f(x)=x^5+ax^4+6x^3+c$有$c\neq 0$，试证：$f(x)$至少有两个复根，它们的虚部不等于零.

16. 设多项式序列$p_1(x)$，$p_2(x)$，…有$p_1(x)=x^2-2$，

$$p_k(x)=p_1(p_{k-1}(x))=p_{k-1}(x)^2-2,\ k=2,3,\cdots$$

试证：对一切自然数n，多项式$p_n(x)-x$的根都是实数，且都是单重根.

17. 给定实系数多项式族$x^4+ax^3+bx^2+ax+1$. 设它们都有实根，试求$\min(a^2+b^2)$.

18. 设$p(x)$为$n-1$次多项式，系数全为正实数. 记a_0为正实数. 试证：n次多项式$f(x)=a_0x^n-p(x)$有唯一的单重正实根.

19. 设$f(x)$为整系数多项式，且$f(0)$和$f(1)$都是奇数，试证：$f(x)$无整数根.

20. 设x^2+px+q和x^2+rx+s都是整系数多项式，且它们有一个公根α不是整数. 试证：$p=r$，$q=s$.

21. 试求整数a,b,c使得$0<|c|<|b|<|a|$，且

$$x(x-a)(x-b)(x-c)+1$$

关于整系数为不可约多项式.

22. 给定n个不同整数$a_1,a_2,\cdots,a_n$，试证：$2n$次整系数多项式$f(x)=\prod_{j=1}^{n}(x-a_j)^2+1$不可约.

23. 设$p(x)$为整系数多项式，且$\deg(p(x))\geqslant 1$. 记适合$p(k)^2=1$的不同整数k的总个数为$n(p)$. 试证：

$$n(p)-\deg(p)\leqslant 2.$$

24. 试求整数a，使得方程组

$$\sum_{k=1}^{5}kx_k=a,\ \sum_{k=1}^{5}k^3x_k=a^2,\ \sum_{k=1}^{5}k^5x_k=a^3$$

有解$x_1^0,\cdots,x_5^0$，其中$x_j^0\geqslant 0$，$1\leqslant j\leqslant 5$. 并求出相应线性方程组的解.

25. 设$f(x)$为有理系数不可约多项式，且$\deg(f(x))=2n+1$，$n\geqslant 1$. 则$f(x)$的任两不同的根之和，差，积，商都不是有理数.

26. 设多项式序列 $f_0(x), f_1(x), \cdots$ 有

$$f_n(2x)=\sum_{j=0}^{n}(-1)^{n-j}\binom{n}{j}2^j f_j(x),\ n=0,1,\cdots$$

试证：对一切数 t，有

$$f_n(tx)=\sum_{j=0}^{n}(-1)^{n-j}\binom{n}{j}t^j(t-1)^{n-j}f_j(x),\quad n=0,1,2,\cdots$$

27. 试求适合条件

$$x^2f(x)+f(1-x)=2x-x^4$$

的多项式 $f(x)$.

28. 试求适合条件

$$xf(x-1)=(x-26)f(x)$$

的多项式 $f(x)$.

29. 给定自然数 n，试求适合条件

$$f(f(x))=f(x)^n$$

的所有多项式.

30. 设 $g(x)$ 为多项式，使得序列 $x_0=0, x_k=g(x_{k-1}), k=1,2,\cdots$，中有无限多个数值不同. 试求多项式 $f(x)$，它适合条件

$$f(g(x))=g(f(x)),\ f(0)=0.$$

引实系数多项式 $f(x)$ 的根都是实根的必要且充分条件为不存在两个非零实系数多项式 $g(x), h(x)$ 使得 $f^2=g^2+h^2$，且 $\deg(g)\neq\deg(h)$.

第二章　行列式理论

§2.1 排　　列

记 $i_1i_2\cdots i_n$ 为 $1,2,\cdots,n$ 的**排列**，它又可记作

$$\begin{pmatrix}1 & 2 & \cdots & n\\ i_1 & i_2 & \cdots & i_n\end{pmatrix}.$$

定义　将 $1,2,\cdots,n$ 的排列 $i_1i_2\cdots i_n$ 变为排列

$$i_1i_2\cdots i_{j-1}i_{j+1}i_ji_{j+2}\cdots i_{n-1}i_n,$$

则称为作了一个**对换**.

显然，对任一排列 $i_1i_2\cdots i_n$，总可以作一系列对换，将它变成**标准排列** $1\,2\cdots n$. 当然，变成标准排列的对换方式不唯一，但对换次数间有关系

引理 2.1.1　给定排列 $i_1i_2\cdots i_n$，任给两种将 $i_1i_2\cdots i_n$ 变为标准排列 $12\cdots n$ 的对换方式，则它们的对换次数之差必为偶数，即对换次数的**奇偶性**只与排列 $i_1i_2\cdots i_n$ 有关.

证　对 n 作归纳法. 当 $n=1$ 时不用讨论. 设对 $n-1$ 个文字 $1,2,\cdots,n-1$ 的排列，引理成立. 下面考虑 n 个文字 $1,2,\cdots,n$ 的排列 $i_1i_2\cdots i_n$. 设 $i_j=n$. 考虑将 $i_1i_2\cdots i_n$ 变为标准排列 $1\,2\cdots n$ 的任一个对换方式，将这个对换方式的所有对换构成一个集合 $\mathfrak{S}$，其中与 $i_j=n$ 无关的对换构成 $\mathfrak{S}$ 的子集合 $\mathfrak{S}_1$，与 $i_j=n$ 有关的对换构成的集合自然为 $\mathfrak{S}_2=\mathfrak{S}-\mathfrak{S}_1$. 显然 $\mathfrak{S}_1$ 构成将 $i_1i_2\cdots i_{j-1}i_{j+1}\cdots i_n$ 变为标准排列 $12\cdots(n-1)$ 的对换方式，由归纳法假设，$\mathfrak{S}_1$ 的元素个数的奇偶性只与排列 $i_1i_2\cdots i_{j-1}i_{j+1}\cdots i_n$，即只与 $i_1i_2\cdots i_n$

有关. 问题化为要证明集合 $\mathfrak{S}_2$ 的元素个数的奇偶性也只与排列 $i_1i_2\cdots i_n$ 有关.

显然 $\mathfrak{S}_2$ 给出将 n 从第 j 个位置经过一系列对换变到第 n 个位置. 这批对换分成两类, 一类是将 n 从第 j 个位置跳到第 $j+1$ 个位置,…,第 n 个位置. 共包含了 $n-j$ 次对换;另一类是由于文字 n 往前跳造成的, 可是只要往前跳 (例如从第 k 个位置跳到第 $k-1$ 个位置) 后, 必然在后面的跳步过程中再往后跳 (例如从第 $k-1$个位置跳到第 k 个位置),否则n 不可能最后跳到第n 个位置. 所以这一类对换必然有偶数个,记作 $2r$. 因此 $\mathfrak{S}_2$ 中元素个数为$n-j+2r$, 它的奇偶性和 $n-j$ 的奇偶性相同. 由于 $i_j=n$, 所以奇偶性只与排列 $i_1i_2\cdots i_n$ 有关. 证完.

由引理 2.1.1,我们可以引进下面定义

定义 若可以经过奇(偶)次对换, 将排列 $i_1i_2\cdots i_n$ 变为标准排列 $1\ 2\ \cdots\ n$, 则排列 $i_1i_2\cdots i_n$ 称为**奇(偶)排列**、这时记

$$\delta^{1\ 2\ \cdots\ n}_{i_1i_2\cdots i_n}=\begin{cases}1, & \text{当 } i_1i_2\cdots i_n \text{ 为偶排列;}\\ -1, & \text{当 } i_1i_2\cdots i_n \text{ 为奇排列.}\end{cases}$$

下面给出排列奇偶性的一系列性质. 由引理 2.1.1 可知, 若存在一种对换方式将排列 $i_1i_2\cdots i_n$ 变为标准排列 $1\ 2\ \cdots\ n$,其对换次数为 s,则 $\delta^{1\ 2\ \cdots\ n}_{i_1i_2\cdots i_n}=(-1)^s$.

引理 2.1.2 n 个文字 $1,2,\ \cdots,\ n$ 的排列 $i_1i_2\cdots i_n$ 和 $i_1i_2\cdots i_{j-1}i_ki_{j+1}\cdots i_{k-1}i_ji_{k+1}\cdots i_n$ 的奇偶性相反,$1\leqslant j<k\leqslant n$.

证 显然, 将排列 $i_1i_2\cdots i_{j-1}i_ki_{j+1}\cdots i_{k-1}i_ji_{k+1}\cdots i_n$ 作 $k-1-j$ 次对换, 可变为排列 $i_1i_2\cdots i_{j-1}i_{j+1}\cdots i_{k-1}i_ki_ji_{k+1}\cdots i_n$. 再作 $k-j$ 次对换, 可变为排列 $i_1i_2\cdots i_{j-1}i_ji_{j+1}\cdots i_{k-1}i_ki_{k+1}\cdots i_n$. 设对 $i_1i_2\cdots i_n$ 作 s 次对换可变为标准排列 $12\cdots n$,于是有这样一种对换方式, 它经过$(k-1-j)+(k-j)+s$ 次对换将排列 $i_1i_2\cdots i_{j-1}i_k i_{j+1}\cdots i_{k-1}i_ji_{k+1}\cdots i_n$ 变为标准排列. 由于 s 和 $s+2(k-j)-$

1 之奇偶性相反，便证明了引理. 证完.

引理 2.1.3 设 $i_1 i_2 \cdots i_n$ 为 n 个文字 $1, 2, \cdots, n$ 的排列，则 $1\,2 \cdots n$ 为 n 个文字 $i_1, i_2, \cdots, i_n$ 的排列，且有

$$\delta^{1\,2\cdots n}_{i_1 i_2 \cdots i_n} = \delta^{i_1 i_2 \cdots i_n}_{1\,2\cdots n}.$$

证 这是因为将 $i_1 i_2 \cdots i_n$ 变为标准排列 $1\,2 \cdots n$ 的对换方式整个逆过来，便得到将 $1\,2 \cdots n$ 变为标准排列 $i_1 i_2 \cdots i_n$ 的对换方式. 所以奇偶性相同. 证完.

引理 2.1.4 设 $i_1 i_2 \cdots i_n$ 和 $j_1 j_2 \cdots j_n$ 都是 $1, 2, \cdots, n$ 的排列，则 $j_1 j_2 \cdots j_n$ 为 $i_1, i_2, \cdots, i_n$ 的排列，且有

$$\delta^{1\,2\cdots n}_{j_1 j_2 \cdots j_n} = \delta^{1\,2\cdots n}_{i_1 i_2 \cdots i_n} \delta^{i_1 i_2 \cdots i_n}_{j_1 j_2 \cdots j_n}.$$

证 由引理 2.1.3. 易知前一结论成立，现在证后一结论. 设 $j_1 j_2 \cdots j_n$ 变为 $i_1 i_2 \cdots i_n$ 的对换方式的次数为 s，再设 $i_1 i_2 \cdots i_n$ 变为 $1\,2 \cdots n$ 的对换方式的次数为 t，合并便成 $j_1 j_2 \cdots j_n$ 变为 $1\,2 \cdots n$ 的对换方式，次数为 $s+t$. 于是

$$\delta^{1\,2\cdots n}_{j_1 j_2 \cdots j_n} = (-1)^{s+t} = (-1)^s (-1)^t = \delta^{1\,2\cdots n}_{i_1 i_2 \cdots i_n} \delta^{i_1 i_2 \cdots i_n}_{j_1 j_2 \cdots j_n}.$$

证完.

引理 2.1.5 设 $j_1 j_2 \cdots j_r$ 为 $1, 2, \cdots, r$ 的排列，$j_{r+1} j_{r+2} \cdots j_n$ 为 $r+1, r+2, \cdots, n$ 的排列，则有

$$\delta^{1\,2\cdots n}_{j_1 j_2 \cdots j_n} = \delta^{1\,2\cdots r}_{j_1 j_2 \cdots j_r} \delta^{r+1\ r+2\cdots n}_{j_{r+1} j_{r+2} \cdots j_n}.$$

证 设 $j_1 j_2 \cdots j_r$ 变为 $12 \cdots r$ 的对换方式，次数为 s，再设 $j_{r+1} j_{r+2} \cdots j_n$ 变为 $(r+1)(r+2) \cdots n$ 的对换方式，次数为 t，合并便成 $j_1 j_2 \cdots j_n$ 变为 $1\,2 \cdots n$ 的对换方式，次数为 $s+t$. 于是

$$\delta^{1\,2\cdots r}_{j_1 j_2 \cdots j_r} \delta^{(r+1)(r+2)\cdots n}_{j_{r+1} j_{r+2} \cdots j_n} = (-1)^{s+t} = \delta^{1\,2\cdots n}_{j_1 j_2 \cdots j_n}.$$

证完.

引理 2.1.6 设 $1, 2, \cdots, n$ 的排列 $i_1 i_2 \cdots i_n$ 有

$$1 \leqslant i_1 < \cdots < i_r \leqslant n, \quad 1 \leqslant i_{r+1} < i_{r+2} < \cdots < i_n \leqslant n$$

则有

$$\delta^{12\cdots n}_{i_1 i_2 \cdots i_n} = (-1)^{i_1+\cdots+i_r+\frac{r(r+1)}{2}}$$

证 作 i_1-1 次对换，可得 $12\cdots n$ 变为 $i_1 12\cdots(i_1-1)(i_1+1)\cdots n$. 今 $i_1<i_2$，再作 i_2-2 次对换，可将它变成

$$i_1 i_2 12\cdots(i_1-1)(i_1+1)\cdots(i_2-1)(i_2+1)\cdots n$$

于是，依次作 i_1-1, i_2-2, $\cdots$, i_r-r 次对换，便将 $12\cdots n$ 变为 $i_1 i_2\cdots i_r j_{r+1}\cdots j_n$. 由于 $1\leqslant j_{r+1}<j_{r+2}<\cdots<j_n\leqslant n$，且 $j_{r+1},\cdots,j_n$ 为由 $1,2,\cdots,n$ 中删去 $i_1,i_2,\cdots,i_r$ 而得到，所以证明了

$$j_k=i_k,\quad r+1\leqslant k\leqslant n.$$

因此证明了

$$\delta^{12\cdots n}_{i_1 i_2\cdots i_n}=(-1)^{(i_1-1)+(i_2-2)+\cdots+(i_r-r)}=(-1)^{i_1+\cdots+i_r+\frac{r(r+1)}{2}}.$$

证完.

引理 2.1.7 设 $1,2,\cdots,n$ 的排列 $i_1i_2\cdots i_n$ 和 $j_1j_2\cdots j_n$ 有

$$1\leqslant i_1<\cdots<i_r\leqslant n,\quad 1\leqslant i_{r+1}<\cdots<i_n\leqslant n,$$

$$1\leqslant j_1<\cdots<j_r\leqslant n,\quad 1\leqslant j_{r+1}<\cdots<j_n\leqslant n,$$

则

$$\delta^{i_1i_2\cdots i_n}_{j_1j_2\cdots j_n}=(-1)^{i_1+\cdots+i_r+j_1+\cdots+j_r}.$$

证 由引理 2.1.4 及引理 2.1.6. 立即可证. 证完.

引理 2.1.8 n 个文字 $1,2,\cdots,n$ 的排列共有 $n!$ 个，当 $n>1$，则奇排列总数和偶排列总数相同，它们都等于 $\frac{1}{2}(n!)$.

证 显然 $1,2,\cdots,n$ 的排列可以用下面方法得到：即取出所有 $1,2,\cdots,n-1$ 的排列 $j_1j_2\cdots j_{n-1}$，于是 $1,2,\cdots、n$ 的所有排列为

$$nj_1j_2\cdots j_{n-1},\ j_1nj_2\cdots j_{n-1},\ \cdots,\ j_1j_2\cdots j_{n-1}n.$$

任取一种来考虑，例如取 $j_1\cdots j_{r-1}nj_r\cdots j_{n-1}$，它经过 $n-r$ 次对换变成排列 $j_1j_2\cdots j_{n-1}n$. 设 $n-r$ 为奇(偶)数，则当 $j_1j_2\cdots j_{n-1}$ 为奇排列时，$j_1\cdots j_{r-1}nj_r\cdots j_{n-1}$ 为偶(奇)排列；当 $j_1j_2\cdots j_{n-1}$ 为偶排列

时，$j_1\cdots j_{r-1}nj_r\cdots j_{n-1}$ 为奇(偶)排列．由归纳法假设，$j_1j_2\cdots j_{n-1}$ 的奇偶排列总个数相同，且等于$\frac{1}{2}(n-1)!$．这证明了 $j_1j_2\cdots j_{r-1}nj_r\cdots j_{n-1}$ 的奇偶排列总个数相同，且等于 $\frac{1}{2}(n-1)!$．于是证明了 n 个文字的排列的奇偶排列的总个数相同，且等于$\frac{1}{2}n!$．所以排列总个数为 $n!$．证完．

在行列式的讨论中，我们还需要用下面和号，例如：

$$\sum_{\binom{12\cdots n}{i_1i_2\cdots i_n}} a_{i_1i_2\cdots i_n}$$

它表示了 $n!$ 个数 $a_{i_1i_2\cdots i_n}$ 之总和，其中 $i_1i_2\cdots i_n$ 取遍 $1,2,\cdots,n$ 的所有排列．

引理 2.1.9 设当指标 $1\leqslant j_1,\cdots,j_p\leqslant n$ 中有两个指标相同，则 $a_{j_1j_2\cdots j_p}=0$，其中 $1\leqslant p\leqslant n$．于是有

$$\sum_{j_1,\cdots,j_p=1}^{n} a_{j_1j_2\cdots j_p}=\sum_{\substack{j_1,\cdots,j_p=1\\ j_1,\cdots,j_p\text{两两不等}}}^{n} a_{j_1j_2\cdots j_p}=\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\ \sum_{\binom{i_1i_2\cdots i_p}{j_1j_2\cdots j_p}} a_{j_1j_2\cdots j_p}$$

证 由条件可知前一等式成立．后一等式证明如下：任取 $1,2,\cdots,n$ 中 p 个不同数 $j_1,j_2,\cdots,j_p$，将它们按大小次序排成 $1\leqslant i_1<i_2<\cdots<i_p\leqslant n$，于是 $j_1j_2\cdots j_p$ 取遍 $i_1,i_2,\cdots,i_p$ 的所有排列，而 $i_1,\cdots,i_p$ 取遍所有适合条件 $1\leqslant i_1<\cdots<i_p\leqslant n$ 的 p 数组．证完．

习题 2.1

记 δ_{ij} 为 Kronecker符号，它定义为：当 $i=j$，则 $\delta_{ij}=1$；当 $i\neq j$，则 $\delta_{ij}=0$. 试证：

$$\sum_{i=1}^{n}a_i\delta_{ij}=a_j,\quad \sum_{i,j=1}^{n}\delta_{ij}=n,\quad \sum_{j=1}^{n}\delta_{ij}\delta_{jk}=\delta_{ik},\quad \sum_{i,j=1}^{n}a_{ij}\delta_{ij}=\sum_{i=1}^{n}a_{ii}.$$

其中 $1\leqslant i,j,k\leqslant n$.

§2.2 n阶行列式

给定 nm 个数 a_{ij}, $1\leqslant i\leqslant n$, $1\leqslant j\leqslant m$, 将它排成矩形

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

称为矩阵，其中横排为行，第 i 行为 $a_{i1}, a_{i2}, \cdots, a_{im}$, $1\leqslant i\leqslant n$; 又竖排为列，第 j 列为

$$\begin{matrix} a_{1j}, \\ a_{2j}, \\ \vdots \\ a_{nj}, \end{matrix}$$

$1\leqslant j\leqslant m$. 所以 A 有 n 个行，m 个列，称为 $n\times m$ 矩阵. 矩阵 A 的第 i 行，第 j 列交叉的元素为 a_{ij}，所以前指标示行，后指标示列. 又 $n\times m$ 矩阵 A 也可记作 $A^{(n,m)}$ 或记作 (a_{ij}).

给定 $n\times m$ 矩阵 A, 取定

$$1\leqslant i_1<\cdots<i_r\leqslant n,\ 1\leqslant j_1<\cdots<j_s\leqslant m,$$

在 A 中取出第 $i_1, i_2, \cdots, i_r$ 行，第 $j_1, j_2, \cdots, j_s$ 列的交叉元素，便构成 $r\times s$ 矩阵

$$\begin{pmatrix} a_{i_1j_1} & \cdots & a_{i_1j_s} \\ \vdots & & \vdots \\ a_{i_rj_1} & \cdots & a_{i_rj_s} \end{pmatrix}$$

称为 $n\times m$ 矩阵 A 的**子矩阵**.

关于矩阵的代数运算是第四章的内容，在这一章，我们只用它的符号.

定义 $n\times n$ 矩阵称为 **n 阶方阵**，n 阶方阵 A 也可记作 $A^{(n)}$.

重要的是引进 n 阶方阵的**行列式**，它是自变量为 n 阶方阵的函数，定义为

定义 n阶方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

的行列式 det A 为表达式

$$\det A=\sum_{\binom{1\ 2\ \cdots\ n}{i_1 i_2 \cdots i_n}} \delta^{1 2 \cdots n}_{i_1 i_2 \cdots i_n} a_{1i_1} a_{2i_2} \cdots a_{ni_n}.$$

det$A$也可记作$|A|$,

由行列式的定义可知:(1) 它由 $n!$ 项构成;(2) 每一项冠以±1,由 $i_1 i_2 \cdots i_n$ 为奇,偶排列所决定;(3) 每一项为$\delta^{1\ 2\ \cdots\ n}_{i_1 i_2 \cdots i_n} a_{1i_1} a_{2i_2} \cdots a_{ni_n}$, 其中 $a_{1i_1} a_{2i_2} \cdots a_{ni_n}$ 即在每一行中取一个且仅取一个元素;同时也是在每一列中取一个且仅取一个元素,再相乘而得. 当$n\geqslant2$, 则恰有 $\frac{1}{2}n!$个项冠以正号, $\frac{1}{2}n!$个项冠以负号.

从定义出发,可以给出行列式的一系列性质. 而计算行列式,又需要充分利用这些性质,由此可见它们的重要性.

定理 2.2.1 行列式有如下性质

(1)
$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}=\det\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix},$$

即给定 n 阶方阵 $A=(a_{ij})$,将它的行变为列,列变为行,使得 A 的第i行,第j列交叉元素 a_{ij} 为第 j 行,第 i 列元素,便得到一个新的n阶方阵,记作 A',称为 A 的**转置方阵**. 于是有 det$A$=detA'.

由 det$A$=detA',所以行列式的有关行的所有性质,对行列式的列也都成立,反之亦然;

(2) 行列式的两行(两列)互换,则行列式变号;

(3) 行列式的一行(一列)乘以常数 λ, 等于行列式本身乘以

常数 λ;

(4) 行列式的两行(两列)成比例，特别若两行(两列)相等，则行列式的值为零. 因此若有一行(一列)全为零，则行列式的值为零.

(5)

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1}+b_{j1} & \cdots & a_{jn}+b_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}=\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

$$+\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_{j1} & \cdots & b_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

对列也有一样的等式.

(6) 将行列式的任一行(列)乘以数 λ，再加到另一行(列)上，则行列式的值不改变.

证 下面分别给出证明.

证 (1). 在行列式定义

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}=\sum\delta_{i_1 i_2\cdots i_n}^{1\,2\cdots n}a_{1i_1}a_{2i_2}\cdots a_{ni_n}$$

中将 $i_1i_2\cdots i_n$ 用一种确定的对换方式变为标准排列 $1\,2\cdots n$，它实际上是由第1个，…，第 n 个位置中的一系列相邻两位置的对换构成. 用这种对换方式作用在标准排列 $1\,2\cdots n$ 上，便变成排列 $j_1j_2\cdots j_n$. 显然，当 $i_1i_2\cdots i_n$ 取遍 $1, 2, \cdots, n$ 的所有排列时，$j_1j_2\cdots j_n$ 也取遍 $1,2,\cdots,n$ 的所有排列，又有

$$\delta^{1\,2\cdots n}_{j_1 j_2 \cdots j_n} = \delta^{j_1 j_2 \cdots j_n}_{1\,2\cdots n} = \delta^{1\,2\cdots n}_{i_1 i_2 \cdots i_n},$$

且有

$$a_{1i_1}a_{2i_2}\cdots a_{ni_n} = a_{j_1 1}a_{j_2 2}\cdots a_{j_n n},$$

这证明了

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{\binom{12\cdots n}{i_1 i_2 \cdots i_n}} \delta^{12\cdots n}_{i_1 i_2 \cdots i_n} a_{1i_1}a_{2i_2}\cdots a_{ni_n}$$

$$= \sum_{\binom{12\cdots n}{j_1 j_2 \cdots j_n}} \delta^{j_1 j_2 \cdots j_n}_{12\cdots n} a_{j_1 1}a\cdots a_{j_n n}$$

$$= \det\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix}.$$

所以证明了(1).

证(2)．取定 $j, k, 1 \leqslant j < k \leqslant n$，则

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{\binom{12\cdots n}{i_1 i_2 \cdots i_n}} \delta^{12\cdots n}_{i_1 i_2 \cdots i_n} a_{1i_1}\cdots a_{ji_k}\cdots a_{ki_j}\cdots a_{ni_n}.$$

由于 $1,2,\cdots,n$ 的排列 $i_1 i_2 \cdots i_n$ 经过 $2(k-j-1)+1$ 次对换，可变为排列 $i_1\cdots i_{j-1} i_k i_{j+1}\cdots i_{k-1} i_j i_{k+1}\cdots i_n$，所以

$$\delta^{12\cdots n}_{i_1 i_2 \cdots i_n} = -\delta^{12\cdots(j-1)j(j+1)\cdots(k-1)k(k+1)\cdots n}_{i_1 i_2 \cdots i_{j-1} i_k i_{j+1}\cdots i_{k-1} i_j i_{k+1}\cdots i_n}.$$

又当 $i_1 i_2 \cdots i_n$ 取遍 $1,2,\cdots,n$ 的所有排列时，$i_1\cdots i_k\cdots i_j\cdots i_n$ 也取遍 $1,2,\cdots,n$ 的所有排列．这证明了

$$\sum_{\binom{12\cdots n}{i_1 i_2 \cdots i_n}} \delta^{12\cdots n}_{i_1 i_2 \cdots i_n} a_{1i_1}a_{2i_2}\cdots a_{ni_n}$$

$$=\sum \delta^{12\cdots(j-1)j(j+1)\cdots(k-1)k(k+1)\cdots n}_{i_1 i_2\cdots i_{j-1} i_k i_{j+1}\cdots i_{k-1} i_j i_{k+1}\cdots i_n} a_{1i_1}a_{2i_2}\cdots a_{j-1,i_{j-1}}a_{ji_k}a_{j+1,i_{j+1}}\cdot$$

$$\cdots a_{k-1,i_{k-1}}a_{ki_j}a_{k+1i,_{k+1}}\cdots a_{ni_n}$$

$$=-\sum_{\binom{12\cdots n}{i_1 i_2\cdots i_n}} \delta^{12\cdots n}_{i_1 i_2\cdots i_n} a_{1i_1}\cdots a_{ji_k}\cdots a_{ki_j}\cdots a_{ni_n}.$$

由行列式的定义，便证明了(2).

证 (3). 记 λ 为常数，则

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{j1} & \cdots & \lambda a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{\binom{12\cdots n}{i_1 i_2\cdots i_n}} \delta^{12\cdots n}_{i_1 i_2\cdots i_n} a_{1i_1}\cdots a_{j-1i_{j-1}},(\lambda a_{ji_j}) \cdot a_{j+1,i_{j+1}}\cdots a_{ni_n}$$

$$=\lambda \sum_{\binom{12\cdots n}{i_1 i_2\cdots i_n}} \delta^{12\cdots n}_{i_1 i_2\cdots i_n} a_{1i_1}\cdots a_{ni_n}.$$

由行列式的定义，便证明了(3).

证 (4). 设行列式

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

的第 j 行 $a_{j1},\cdots,a_{jn}$ 和第 k 行 $a_{k1},\cdots,a_{kn}$ 成比例，其中 $j\neq k$. 于是存在常数 λ，使得

$$a_{ki}=\lambda a_{ji},\quad i=1,2,\cdots,n.$$

由性质(3)，可提出常数因子 λ，即有

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \lambda\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{matrix} \\ \\ (j) \\ \\ (k) \\ \\ \\ \end{matrix}$$

后一行列式的第 j 行和第 k 行互换，由于第 j 行和第 k 行相同，所以行列式的值不改变. 但是由性质(2)，行列式的值变号. 这证明了后一行列式必须为零，所以前一行列式的值为零. 这证明了(4).

证(5). 由行列式的定义，

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1}+b_{j1} & \cdots & a_{jn}+b_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{\binom{12\cdots n}{i_1 i_2 \cdots i_n}} \delta^{12\cdots n}_{i_1 i_2 \cdots i_n} a_{1i_1}\cdots(a_{ji_j}+b_{ji_j})\cdots a_{ni_n}$$

$$= \sum_{\binom{12\cdots n}{i_1 i_2 \cdots i_n}} \delta^{12\cdots n}_{i_1 i_2 \cdots i_n} a_{1i_1}\cdots a_{ji_j}\cdots a_{ni_n} + \sum_{\binom{12\cdots n}{i_1 i_2 \cdots i_n}} \delta^{12\cdots n}_{i_1 i_2 \cdots i_n} a_{1i_1}\cdots b_{ji_j}\cdots a_{ni_n}$$

由行列式的定义，便证明了(5).

证(6). 由性质(5)，取 $j\neq k, 1\leqslant j,k\leqslant n$,

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1}+\lambda a_{k1} & \cdots & a_{jn}+\lambda a_{kn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}(j) = \det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}(j)$$

$$+\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{k1} & \cdots & \lambda a_{kn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}(j)$$

由于 $k\neq j$，由性质(4)，便证明了性质(6). 定理证完.

习题 2.2

1. 试从定义出发来证明恒等式:

$$\det\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = a_{11}\det\begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

由此推出

$$\det\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

2. 试用行列式的性质来证明:

$$\det\begin{pmatrix} b+c & c+a & a+b \\ b_1+c_1 & c_1+a_1 & a_1+b_1 \\ b_2+c_2 & c_2+a_2 & a_2+b_2 \end{pmatrix} = 2\det\begin{pmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix},$$

$$\det\begin{pmatrix} 1 & a & a^2-bc \\ 1 & b & b^2-ca \\ 1 & c & c^2-ab \end{pmatrix} = 0.$$

3. 试计算下列行列式:

$$\det\begin{pmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{pmatrix},$$

$$\det\begin{pmatrix} \cos\theta & \sin\theta & \cos\theta & \sin\theta \\ \cos2\theta & \sin2\theta & 2\cos2\theta & 2\sin2\theta \\ \cos3\theta & \sin3\theta & 3\cos3\theta & 3\sin3\theta \\ \cos4\theta & \sin4\theta & 4\cos4\theta & 4\sin4\theta \end{pmatrix}.$$

4. 试证下列等式:

$$\det\begin{pmatrix} 1 & \mathrm{tg}A & \sin2A \\ 1 & \mathrm{tg}B & \sin2B \\ 1 & \mathrm{tg}C & \sin2C \end{pmatrix} = 0, A+B+C=\pi;$$

$$\det\begin{pmatrix} \cos(a-b) & \cos(a+b) & \sin(a+b) \\ \cos(b-c) & \cos(b+c) & \sin(b+c) \\ \cos(c-a) & \cos(c+a) & \sin(c+a) \end{pmatrix} =$$

$$-2\sin(a-b)\sin(b-c)\sin(c-a).$$

5. 给定 n^2 个数，将它们任意排成一个 n 阶方阵，试证：它们的行列式至多有 $(n^2)!/(n!)^2$ 个不同的值.

6. 试证：对 n 阶方阵 (a_{ij})，

$$\det(a_{ij})=\det((-1)^{i+j}a_{ij}).$$

§2.3 代数余子式和 Laplace 展开

取定 n 阶方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

任取 $1\leqslant i_1<\cdots<i_p\leqslant n, 1\leqslant j_1<\cdots<j_p\leqslant n$. 记 p 阶行列式

$$A\begin{pmatrix} i_1 & i_2\cdots i_p \\ j_1 & j_2\cdots j_p \end{pmatrix}=\det\begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_p} \\ \vdots & & \vdots \\ a_{i_p j_1} & \cdots & a_{i_p j_p} \end{pmatrix},$$

则 $A\begin{pmatrix} i_1 & i_2\cdots i_p \\ j_1 & j_2\cdots j_p \end{pmatrix}$称为行列式 $\det A$ 的 **p 阶子式**. 特别，$A\begin{pmatrix} i_1 i_2\cdots i_p \\ i_1 i_2\cdots i_p \end{pmatrix}$称为 **$p$ 阶主子式**.

在 $1,2,\cdots,n$ 中除去 $i_1, i_2, \cdots, i_p$, 再按大小次序排为 $1\leqslant i_{p+1}<\cdots<i_n\leqslant n$, 在 $1,2,\cdots,n$ 中除去 $j_1,j_2,\cdots, j_p$, 再按大小次序排为 $1\leqslant j_{p+1}<\cdots<j_n\leqslant n$. 则

$$(-1)^{i_1+\cdots+i_p+j_1+\cdots+j_p}\det\begin{pmatrix} a_{i_{p+1}j_{p+1}} & \cdots & a_{i_{p+1}j_n} \\ \vdots & & \vdots \\ a_{i_n j_{p+1}} & \cdots & a_{i_n j_n} \end{pmatrix}$$

称为 p 阶子式 $A\begin{pmatrix} i_1\cdots i_p \\ j_1\cdots j_p \end{pmatrix}$的**代数余子式**.

由引理 2.1.7，$A\begin{pmatrix} i_1\cdots i_p \\ j_1\cdots j_p \end{pmatrix}$的代数余子式也可写为：

$$\delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n} A\begin{pmatrix} i_{p+1}\cdots i_n \\ j_{p+1}\cdots j_n \end{pmatrix}.$$

证明 Laplace 展开式的关键是证明

引理 2.3.1 在 $1,2,\cdots,n$ 中取定自然数 p，则有

$$\sum_{\binom{1\,2\cdots n}{i_1 i_2\cdots i_n}} a_{i_1 i_2\cdots i_n}=\sum_{1\leqslant j_1<\cdots<j_p\leqslant n}\ \sum_{\binom{j_1 j_2\cdots j_p}{i_1 i_2\cdots i_p}}\ \sum_{\binom{j_{p+1} j_{p+2}\cdots j_n}{i_{p+1} i_{p+2}\cdots i_n}} a_{i_1 i_2\cdots i_n},$$

其中 $1\leqslant j_{p+1}<\cdots<j_n\leqslant n$，且 $j_1j_2\cdots j_pj_{p+1}j_{p+2}\cdots j_n$ 为 1, 2, …, n 之排列.

证 显然后一和式中每项 $a_{i_1i_2\cdots i_n}$ 中的 $i_1i_2\cdots i_n$ 为 $1,2,\cdots,n$ 之排列，且和式总共有 $\binom{n}{p}\cdot p!\,(n-p)!=n!$ 项，前一和式总共也有 $n!$ 项. 所以为了证明引理，只要证明后一和式中出现的两组指标 $k_1k_2\cdots k_n$ 和 $l_1l_2\cdots l_n$ 为 1, 2, …, n 的两个不同的排列. 今若 $k_1k_2\cdots k_p$ 和 $l_1l_2\cdots l_p$ 为两个不同的排列，自然 $k_1k_2\cdots k_n$ 和 $l_1l_2\cdots l_n$ 也为两个不同的排列. 若 $k_1k_2\cdots k_p$ 和 $l_1l_2\cdots l_p$ 为两个相同的排列，于是 $l_1=k_1,\cdots,l_p=k_p$. 将 $k_1,k_2,\ \cdots,\ k_p$ 按大小次序排成 $1\leqslant j_1<\cdots<j_p\leqslant n$，于是 $k_1k_2\cdots k_p=l_1l_2\cdots l_p$ 为 $j_1,j_2,\cdots,j_p$ 的排列. 它们对应的项自然地落在

$$\sum_{\binom{j_1 j_2\cdots j_p}{i_1 i_2\cdots i_p}}\ \sum_{\binom{j_{p+1} j_{p+2}\cdots j_n}{i_{p+1} i_{p+2}\cdots i_n}} a_{i_1 i_2\cdots i_p}$$

中，即在

$$\sum_{\binom{j_{p+1} j_{p+2}\cdots j_n}{i_{p+1} i_{p+2}\cdots i_n}} a_{k_1k_2\cdots k_p i_{p+1} i_{p+2}\cdots i_n}$$

中，其中 $j_{p+1}<j_{p+2}<\cdots<j_n$ 是由 $1,2,\cdots,n$ 除去 $j_1,j_2,\cdots,j_p$ 后所得到的数集. 由上面和式的定义可知 $a_{k_1k_2\cdots k_n}$ 和 $a_{l_1l_2\cdots l_n}$ 的指标中，$k_{p+1}k_{p+2}\cdots k_n$ 及 $l_{p+1}l_{p+2}\cdots l_n$ 为 $j_{p+1},j_{p+2},\cdots,j_n$ 的不同排列，因此 $k_1k_2\cdots k_n$ 及 $l_1l_2\cdots l_n$ 仍为 $1,2,\cdots,n$ 的不同排列. 证完.

引理 2.3.2 n 阶行列式 $\det(a_{ij})$ 的子式

$$A\binom{i_1 i_2\cdots i_p}{j_1 j_2\cdots j_p}=\sum_{\binom{j_1 j_2\cdots j_p}{l_1 l_2\cdots l_p}}\delta^{j_1 j_2\cdots j_p}_{l_1 l_2\cdots l_p} a_{i_1 l_1} a_{i_2 l_2}\cdots a_{i_p l_p}.$$

证 由行列式及子式的定义立即可证. 证完.

定理 2.3.1(Laplace 展开定理) 任取指标

$$1\leqslant i_1<i_2<\cdots<i_p\leqslant n,$$

则 n 阶行列式 $\det(a_{ij})$ 可按照第 $i_1, i_2, \cdots, i_p$ 行作 Laplace 展开

$$\det(a_{ij})=\sum_{1\leqslant j_1<\cdots<j_p\leqslant n} A\binom{i_1 i_2\cdots i_p}{j_1 j_2\cdots j_p}\left[\delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n} A\binom{i_{p+1} i_{p+2}\cdots i_n}{j_{p+1} j_{p+2}\cdots j_n}\right],$$

其中 $1\leqslant i_{p+1}<\cdots<i_n\leqslant n$, $1\leqslant j_{p+1}<\cdots<j_n\leqslant n$, 且 $i_1i_2\cdots i_n$ 和 $j_1j_2\cdots j_n$ 都是 $1,2,\cdots,n$ 的排列. 又可按照第 $i_1, i_2,\cdots, i_p$ 列作 Laplace 展开

$$\det(a_{ij})=\sum_{1\leqslant j_1<\cdots<j_p\leqslant n} A\binom{j_1 j_2\cdots j_p}{i_1 i_2\cdots i_p}\left[\delta^{j_1 j_2\cdots j_n}_{i_1 i_2\cdots i_n} A\binom{j_{p+1} j_{p+2}\cdots j_n}{i_{p+1} i_{p+2}\cdots i_n}\right],$$

指标说明同上.

证 由于转置方阵的行列式等于原方阵之行列式. 所以只要证明前一等式即可. 由定理 2.2.1 的性质(2),

$$\det\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn}\end{pmatrix}=(-1)^{i_1+\cdots+i_p+\frac{p(p+1)}{2}}\det\begin{pmatrix} a_{i_1 1} & \cdots & a_{i_1 n} \\ \vdots & & \vdots \\ a_{i_n 1} & \cdots & a_{i_n n}\end{pmatrix}$$

$$=\delta^{1\,2\cdots n}_{i_1 i_2\cdots i_n}\sum_{\binom{1\,2\cdots n}{t_1 t_2\cdots t_n}}\delta^{1\,2\cdots n}_{t_1 t_2\cdots t_n} a_{i_1 t_1}a_{i_2 t_2}\cdots a_{i_n t_n}$$

$$=\delta^{1\,2\cdots n}_{i_1 i_2\cdots i_n}\sum_{1\leqslant j_1<\cdots<j_p\leqslant n}\ \sum_{\binom{j_1\cdots j_p}{t_1\cdots t_p}}\ \sum_{\binom{j_{p+1}\cdots j_n}{t_{p+1}\cdots t_n}}\delta^{1\,2\cdots n}_{t_1 t_2\cdots t_n} a_{i_1 t_1}a_{i_2 t_2}\cdots a_{i_n t_n}$$

$$=\sum_{1\leqslant j_1<\cdots<j_p\leqslant n}\delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n}\sum_{\binom{j_1\cdots j_p}{t_1\cdots t_p}}\ \sum_{\binom{j_{p+1}\cdots j_n}{t_{p+1}\cdots t_n}}\delta^{j_1 j_2\cdots j_n}_{t_1 t_2\cdots t_n} a_{i_1 t_1}a_{i_2 t_2}\cdots a_{i_n t_n}$$

$$=\sum_{1\leqslant j_1<\cdots<j_p\leqslant n} A\binom{i_1 i_2\cdots i_p}{j_1 j_2\cdots j_p}\left[\delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n} A\binom{i_{p+1} i_{p+2}\cdots i_n}{j_{p+1} j_{p+2}\cdots j_n}\right].$$

证完.

下面有两个重要推论,

推论 1

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 0 \\ a_{n+1,1} & \cdots & a_{n+1,n} & a_{n+1,n+1} & \cdots & a_{n+1,n+m} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n+m,1} & \cdots & a_{n+m,n} & a_{n+m,n+1} & \cdots & a_{n+m,n+m} \end{pmatrix}$$

$$=\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\det\begin{pmatrix} a_{n+1,n+1} & \cdots & a_{n+1,n+m} \\ \vdots & & \vdots \\ a_{n+m,n+1} & \cdots & a_{n+m,n+m} \end{pmatrix}$$

推论 2 (关于一行及一列的 Laplace 展开)

$$\sum_{j=1}^{n} a_{ij}A_{kj}=\delta_{ik}\det A,$$

$$\sum_{j=1}^{n} a_{ji}A_{jk}=\delta_{ik}\det A,$$

其中

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad A_{pq}=(-1)^{p+q}A\begin{pmatrix} 1\,2\cdots(p-1)(p+1)\cdots n \\ 1\,2\cdots(q-1)(q+1)\cdots n \end{pmatrix}.$$

证 推论 1 用关于前 n 行的 Laplace 展开立即可证. 后一推论用关于 1 行(及 1 列)的 Laplace 展开, 例如:

$$\sum_{j=1}^{n} a_{ij}A_{kj}=\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}(k),$$

当 $k=i$ 有 $\sum_{j=1}^{n} a_{ij}A_{ij}=\det A$; 当$k\neq i$, 由定理 2.2.1 之性质(4),

所以$\sum_{j=1}^{n} a_{ij}A_{kj}=0$. 证完.

利用推论 2,还可以给出行列式的等价定义

定义 n 阶方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

的行列式 $\det A$ 是 A 的元素 $a_{11},\cdots,a_{1n},\cdots,a_{n1},\cdots,a_{nn}$ 的函数,它按归纳法定义为

(i) 当$n=1$, $\det(a_{11})=a_{11}$,

(ii) 当 $n>1$,假设已定义好所有 $n-1$ 阶方阵的行列式,则 n 阶行列式定义为

$$\det A=a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n}.$$

注意,这两种行列式定义是等价的. 由前一种定义可推出后一种定义,这是用 Laplace展开. 而由后一种定义可推出前一种定义,这可在第八章中找到.

习题 2.3

1. 求行列式

$$\det\begin{pmatrix} 0 & 0 & 0 & a_1 \\ 0 & 0 & a_2 & 0 \\ 0 & a_3 & 0 & 0 \\ a_4 & 0 & 0 & 0 \end{pmatrix}$$

的一切元素的三阶代数余子式之和.

2. 给定 n^2 个关于自变量 t 的可微函数 $a_{ij}(t)$, $1\leqslant i,j\leqslant n$. 作 n 阶方阵

$$A(t)=\begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix}.$$

试证

$$\frac{d}{dt}(\det A(t))=\sum_{i,j=1}^{n}\frac{da_{ij}(t)}{dt}A_{ij},$$

其中 A_{ij} 为 n 阶方阵 $A(t)$ 的元素 a_{ij} 的代数余子式，$1\leqslant i,j\leqslant n$.

3. 将 n 阶方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} & 1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & 1 \end{pmatrix}$$

的第 j 行易为 $x_1, x_2, \cdots, x_{n-1}$, 1，得到的新的 n 阶方阵记作 A_j，$1\leqslant j\leqslant n$.
试证:

$$\det A=\sum_{j=1}^{n}\det A_j$$

4. 设 n 阶方阵 $A=(a_{ij})$ 有

$$\sum_{j=1}^{n}a_{ij}=\sum_{j=1}^{n}a_{ji}=0,\ i=1,2,\cdots,n.$$

试证: n 阶方阵 A 的 n^2 个代数余子式 A_{ij} 的值都相等.

5. 记

$$b_{ij}=a_{i1}+\cdots+a_{in}-aa_{ij},\ 1\leqslant i,\ j\leqslant n,$$

试证:

$$\det\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}=(-a)^{n-1}(n-a)\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

§2.4 行列式计算的一些技巧

例 1 Vandermonde行列式

$$V_n(a_1,\cdots,a_n)=\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{pmatrix}=\prod_{1\leqslant i<j\leqslant n}(a_j-a_i).$$

证 将第 $n-1$ 行乘以 $-a_1$ 然后加到第 n 行上，再将第 $n-2$ 行

乘以 $-a_1$ 然后加到第 $n-1$ 行上,…,再将第1行乘以 $-a_1$ 然后加到第2行上,于是有

$$V_n(a_1,\cdots,a_n)=\det\begin{pmatrix}1 & 1 & \cdots & 1\\ 0 & a_2-a_1 & \cdots & a_n-a_1\\ 0 & a_2(a_2-a_1) & \cdots & a_n(a_n-a_1)\\ \vdots & \vdots & & \vdots\\ 0 & a_2^{n-2}(a_2-a_1) & \cdots & a_n^{n-2}(a_n-a_1)\end{pmatrix}$$

$$=\prod_{j=2}^{n}(a_j-a_1)\det\begin{pmatrix}1 & \cdots & 1\\ a_2 & \cdots & a_n\\ \vdots & & \vdots\\ a_2^{n-2} & \cdots & a_n^{n-2}\end{pmatrix}$$

$$=\prod_{j=2}^{n}(a_j-a_1)V_{n-1}(a_2,\cdots,a_n).$$

利用归纳法便证明了例1. 证完.

例2 计算三对角行列式

$$A_n=\det\begin{pmatrix}a & b & & & & \\ c & a & b & & & \\ & c & a & \ddots & b & \\ & & c & \ddots & & \\ & & & \ddots & a & b\\ & & & & c & a\end{pmatrix}$$

其中未写出的元素全是零.

解 对第1行作 Laplace 展开,则有

$$A_n=aA_{n-1}-bcA_{n-2},\ n=3,4,\cdots$$

而

$$A_1=a,\quad A_2=a^2-bc.$$

问题化为求线性递推公式

$$x_n=\lambda x_{n-1}+\mu x_{n-2},\quad n=3,4,\cdots$$

在给定初值 $x_1=\sigma, x_2=\tau$ 时的解. 作法如下:

第一步. 求 α, β 使得

$$\alpha+\beta=\lambda, \quad \alpha\beta=-\mu.$$

于是递推公式变为

$$x_n=(\alpha+\beta)x_{n-1}-\alpha\beta x_{n-2}, \quad n=3,4,\cdots$$

第二步 令

$$y_n=x_n-\alpha x_{n-1}, \quad z_n=x_n-\beta x_{n-1}, \quad n=3,4,\cdots$$

于是有

$$y_n=\beta y_{n-1}, \quad z_n=\alpha z_{n-1}, \quad n=3,4,\cdots$$

因此

$$y_n=\beta^{n-2}y_2, \quad z_n=\alpha^{n-2}z_2$$

其中

$$y_2=x_2-\alpha x_1=\tau-\alpha\sigma, \quad z_2=x_2-\beta x_1=\tau-\beta\sigma.$$

代回,有

$$x_n-\alpha x_{n-1}=\beta^{n-2}(\tau-\alpha\sigma), \quad n=3,4,\cdots$$

$$x_n-\beta x_{n-1}=\alpha^{n-2}(\tau-\beta\sigma), \quad n=3,4,\cdots$$

当 $\alpha\neq\beta$,有

$$(\alpha-\beta)x_{n-1}=\tau(\alpha^{n-2}-\beta^{n-2})+\alpha\beta\sigma(\beta^{n-3}-\alpha^{n-3}), \quad n=3,4,\cdots$$

于是

$$x_{n-1}=\tau\frac{\alpha^{n-2}-\beta^{n-2}}{\alpha-\beta}-\alpha\beta\sigma\frac{\alpha^{n-3}-\beta^{n-3}}{\alpha-\beta}, \quad n=3,4,\cdots$$

所以解为

$$x_n=\tau\frac{\alpha^{n-1}-\beta^{n-1}}{\alpha-\beta}-\alpha\beta\sigma\frac{\alpha^{n-2}-\beta^{n-2}}{\alpha-\beta}, \quad n=3,4,\cdots$$

代回原式验证为恒等式.

当 $\alpha=\beta$,有

$$x_n=\alpha x_{n-1}+\alpha^{n-2}\tau-\alpha^{n-1}\sigma, \quad n=3,4,\cdots$$

于是有

$$x_3=\alpha x_2+\alpha\tau-\alpha^2\sigma,$$
$$x_4=\alpha x_3+\alpha^2\tau-\alpha^3\sigma,$$
$$\cdots\cdots$$
$$x_{n-1}=\alpha x_{n-2}+\alpha^{n-3}\tau-\alpha^{n-2}\sigma,$$
$$x_n=\alpha x_{n-1}+\alpha^{n-2}\tau-\alpha^{n-1}\sigma.$$

依次乘 $\alpha^{n-3},\alpha^{n-4},\cdots,\alpha,1$，再全体相加，则有

$$\begin{aligned}x_n&=\alpha^{n-2}x_2+(n-2)(\alpha^{n-2}\tau-\alpha^{n-1}\sigma)\\&=(n-1)\alpha^{n-2}\tau-(n-2)\alpha^{n-1}\sigma,\quad n=3,4,\cdots\end{aligned}$$

代回原式验证为恒等式.

令 $\lambda=a,\mu=-bc,\sigma=a,\tau=a^2-bc$ 代回便求出行列式 A_n 的值. 由于结论复杂，就不再写出.

例 3 求证：

$$\det\begin{pmatrix}a_{11}+x_1 & \cdots & a_{1n}+x_n\\ \vdots & & \vdots\\ a_{n1}+x_1 & \cdots & a_{nn}+x_n\end{pmatrix}=\det\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}+\sum_{j=1}^{n}x_j\left(\sum_{i=1}^{n}A_{ij}\right)$$

证

$$\det\begin{pmatrix}a_{11}+x_1 & \cdots & a_{1n}+x_n\\ \vdots & & \vdots\\ a_{n1}+x_1 & \cdots & a_{nn}+x_n\end{pmatrix}=\det\begin{pmatrix}a_{11} & a_{12}+x_2 & \cdots & a_{1n}+x_n\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2}+x_2 & \cdots & a_{nn}+x_n\end{pmatrix}$$

$$+\det\begin{pmatrix}x_1 & a_{12}+x_2 & \cdots & a_{1n}+x_n\\ \vdots & \vdots & & \vdots\\ x_1 & a_{n2}+x_2 & \cdots & a_{nn}+x_n\end{pmatrix}=\cdots=\det\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}$$

$$+\sum_{j=1}^{n}\det\begin{pmatrix}a_{11} & \cdots & a_{1,j-1} & x_j & a_{1,j+1} & \cdots & a_{1n}\\ \vdots & & \vdots & \vdots & \vdots & & \vdots\\ a_{n1} & \cdots & a_{n,j-1} & x_j & a_{n,j+1} & \cdots & a_{nn}\end{pmatrix}.$$

用 Laplace 展开，便证明了例 3. 证完.

例 4　试证：

$$d_n=\det\begin{pmatrix} 1 & \cdots & 1 \\ a_{21}-a_{11} & \cdots & a_{2n}-a_{1n} \\ \vdots & & \vdots \\ a_{n1}-a_{n-1,1} & \cdots & a_{nn}-a_{n-1,n} \end{pmatrix}=\sum_{i,j=1}^{n} A_{ij}.$$

证　将行列式 d_n 的第 $1,2,\cdots,n-1$ 行全加到第 n 行上，再将第 $1,2,\cdots,n-2$ 行全加到第 $n-1$ 行上，$\cdots$，便得到

$$d_n=\det\begin{pmatrix} a_{11}+(1-a_{11}) & \cdots & a_{1n}+(1-a_{1n}) \\ a_{21}+(1-a_{11}) & \cdots & a_{2n}+(1-a_{1n}) \\ \vdots & & \vdots \\ a_{n1}+(1-a_{11}) & \cdots & a_{nn}+(1-a_{1n}) \end{pmatrix}.$$

由例 3 可知

$$d_n=\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}+\sum_{j=1}^{n}(1-a_{1j})\ \sum_{i=1}^{n}A_{ij}$$

由 Laplace 展开定理，有

$$d_n=\sum_{j=1}^{n}a_{1j}A_{1j}+\sum_{i,j=1}^{n}A_{ij}-\sum_{i=1}^{n}\left(\sum_{j=1}^{n}a_{1j}A_{ij}\right)=\sum_{i,j=1}^{n}A_{ij}.$$

证完.

例 5　设 $a_1,a_2,\cdots,a_n$ 都不等于零，试求

$$d_n=\det\begin{pmatrix} 0 & a_1+a_2 & \cdots & a_1+a_n \\ a_2+a_1 & 0 & \cdots & a_2+a_n \\ \vdots & \vdots & & \vdots \\ a_n+a_1 & a_n+a_2 & \cdots & 0 \end{pmatrix}$$

解　将它变成 $n+1$ 阶行列式

$$d_n=\det\begin{pmatrix}1 & a_1 & a_2 & \cdots & a_n\\ 0 & 0 & a_1+a_2 & \cdots & a_1+a_n\\ 0 & a_2+a_1 & 0 & \cdots & a_2+a_n\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & a_n+a_1 & a_n+a_2 & \cdots & 0\end{pmatrix}.$$

再将第 2,3,…,$n+1$ 行分别减去第 1 行,便得

$$d_n=\det\begin{pmatrix}1 & a_1 & a_2 & \cdots & a_n\\ -1 & -a_1 & a_1 & \cdots & a_1\\ -1 & a_2 & -a_2 & \cdots & a_2\\ \vdots & \vdots & \vdots & & \vdots\\ -1 & a_n & a_n & \cdots & -a_n\end{pmatrix}.$$

再将它变成 $n+2$ 阶行列式

$$d_n=\det\begin{pmatrix}1 & 0 & 0 & 0 & \cdots & 0\\ 0 & 1 & a_1 & a_2 & \cdots & a_n\\ a_1 & -1 & -a_1 & a_1 & \cdots & a_1\\ a_2 & -1 & a_2 & -a_2 & \cdots & a_2\\ \vdots & \vdots & \vdots & \vdots & & \vdots\\ a_n & -1 & a_n & a_n & \cdots & -a_n\end{pmatrix}.$$

将第 3,4,…,$n+2$ 列分别减去第 1 列，便得

$$d_n=\det\begin{pmatrix}1 & 0 & -1 & -1 & \cdots & -1\\ 0 & 1 & a_1 & a_2 & \cdots & a_n\\ a_1 & -1 & -2a_1 & 0 & \cdots & 0\\ a_2 & -1 & 0 & -2a_2 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & & \vdots\\ a_n & -1 & 0 & 0 & \cdots & -2a_n\end{pmatrix}.$$

再将第 3,4，…，$n+2$ 列分别乘以$\frac{1}{2}$再一起加到第 1 列上，将第 3,4,…,$n+2$列分别乘以$-\frac{1}{2a_1},\cdots,-\frac{1}{2a_n}$再一起加到第 2 列上，于

是有

$$d_n=\det\begin{pmatrix} 1-\frac{n}{2} & \frac{1}{2}\sum_{j=1}^{n}\frac{1}{a_j} & -1 & -1 & \cdots & -1 \\ \frac{1}{2}\sum_{j=1}^{n}a_j & 1-\frac{n}{2} & a_1 & a_2 & \cdots & a_n \\ 0 & 0 & -2a_1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -2a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -2a_n \end{pmatrix}$$

$$=\left[\det\begin{pmatrix} 1-\frac{n}{2} & \frac{1}{2}\sum_{j=1}^{n}\frac{1}{a_j} \\ \frac{1}{2}\sum_{j=1}^{n}a_j & 1-\frac{n}{2} \end{pmatrix}\right](-2)^n\prod_{j=1}^{n}a_j$$

$$=(-2)^n\left(\prod_{j=1}^{n}a_j\right)\left[\left(1-\frac{n}{2}\right)^2-\frac{1}{4}\sum_{i,k=1}^{n}\frac{a_j}{a_k}\right].$$

例6 求下面 n 阶行列式

$$d_n=\det\begin{pmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{pmatrix}.$$

下面给出六种解法，有些是大同小异的.

解1 将它变成 $n+1$ 阶行列式，即有

$$d_n=\det\begin{pmatrix} 1 & a & a & \cdots & a \\ 0 & x & a & \cdots & a \\ 0 & a & x & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a & a & \cdots & x \end{pmatrix}$$

将第 2, 3, …, $n+1$ 行减去第 1 行，便得

$$d_n=\det\begin{pmatrix}1 & a & a & \cdots & a\\ -1 & x-a & 0 & \cdots & 0\\ -1 & 0 & x-a & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ -1 & 0 & 0 & \cdots & x-a\end{pmatrix}.$$

视 x 为符号，将第 2, 3, …, $n+1$ 列分别乘以 $\frac{1}{x-a}$ 再一起加到第 1 列上，便有

$$d_n=\det\begin{pmatrix}1+\frac{na}{x-a} & a & a & \cdots & a\\ 0 & x-a & 0 & \cdots & 0\\ 0 & 0 & x-a & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & x-a\end{pmatrix}$$

$$=\left(1+\frac{na}{x-a}\right)(x-a)^n=(x-a)^n+na(x-a)^{n-1}$$

$$=[x+(n-1)a](x-a)^{n-1}.$$

解 2 将 d_n 的第 2, 3, …, n 列全部加到第 1 列上，便有

$$d_n=\det\begin{pmatrix}(n-1)a+x & a & \cdots & a\\ (n-1)a+x & x & \cdots & a\\ \vdots & \vdots & & \vdots\\ (n-1)a+x & a & \cdots & x\end{pmatrix}$$

$$=[(n-1)a+x]\det\begin{pmatrix}1 & a & \cdots & a\\ 1 & x & \cdots & a\\ \vdots & \vdots & & \vdots\\ 1 & a & \cdots & x\end{pmatrix}$$

再将第 2, 3, …, n 行分别减去第 1 行，便有

$$d_n=[(n-1)a+x]\det\begin{pmatrix}1 & a & \cdots & a\\ 0 & x-a & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & x-a\end{pmatrix}$$

$$=[(n-1)a+x](x-a)^{n-1}.$$

解 3 将第 2, 3, …, n 行分别减去第 1 行，便有

$$d_n=\det\begin{pmatrix}x & a & \cdots & a\\ a-x & x-a & \cdots & 0\\ \vdots & \vdots & & \vdots\\ a-x & 0 & \cdots & x-a\end{pmatrix}$$

再将第 2, 3, …, n 列全都加到第 1 列上去，便有

$$d_n=\det\begin{pmatrix}x+(n-1)a & a & \cdots & a\\ 0 & x-a & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & x-a\end{pmatrix}$$

$$=[x+(n-1)a](x-a)^{n-1}.$$

解 4

$$d_n=\det\begin{pmatrix}a+(x-a) & a & \cdots & a\\ a & x & \cdots & a\\ \vdots & \vdots & & \vdots\\ a & a & \cdots & x\end{pmatrix}$$

$$=\det\begin{pmatrix}a & a & \cdots & a\\ a & x & \cdots & a\\ \vdots & \vdots & & \vdots\\ a & a & \cdots & x\end{pmatrix}+\det\begin{pmatrix}x-a & a & \cdots & a\\ 0 & x & \cdots & a\\ \vdots & \vdots & & \vdots\\ 0 & a & \cdots & x\end{pmatrix}$$

$$=\det\begin{pmatrix}a & 0 & \cdots & 0\\ a & x-a & \cdots & 0\\ \vdots & \vdots & & \vdots\\ a & 0 & \cdots & x-a\end{pmatrix}+(x-a)d_{n-1}$$

$$=a(x-a)^{n-1}+(x-a)d_{n-1}.$$

于是有递推公式

$$d_n=(x-a)d_{n-1}+a(x-a)^{n-1},\quad d_1=x$$

详细写为

$$\begin{aligned}
d_2&=(x-a)d_1+a(x-a),\\
d_3&=(x-a)d_2+a(x-a)^2,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
d_{n-1}&=(x-a)d_{n-2}+a(x-a)^{n-2},\\
d_n&=(x-a)d_{n-1}+a(x-a)^{n-1}.
\end{aligned}$$

依次乘$(x-a)^{n-2}$, $(x-a)^{n-3}$, $\cdots$, $x-a$, 1, 便有

$$\begin{aligned}
(x-a)^{n-2}d_2&=(x-a)^{n-1}d_1+a(x-a)^{n-1},\\
(x-a)^{n-3}d_3&=(x-a)^{n-2}d_2+a(x-a)^{n-1},\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
(x-a)d_{n-1}&=(x-a)^2d_{n-2}+a(x-a)^{n-1},\\
d_n&=(x-a)d_{n-1}+a(x-a)^{n-1}.
\end{aligned}$$

全体加起来，便有

$$\begin{aligned}
d_n&=(x-a)^{n-1}d_1+(n-1)a(x-a)^{n-1}\\
&=[x+(n-1)a](x-a)^{n-1}.
\end{aligned}$$

解5 对行列式d_n的第1行作Laplace展开,有

$$d_n=xd_{n-1}+a\sum_{l=1}^{n-1}(-1)^l\det\begin{pmatrix}
a & x & \cdots & a & a & \cdots & a\\
\vdots & \vdots & & \vdots & \vdots & & \vdots\\
a & a & \cdots & x & a & \cdots & a\\
a & a & \cdots & a & a & \cdots & a\\
a & a & \cdots & a & x & \cdots & a\\
\vdots & \vdots & & \vdots & \vdots & & \vdots\\
a & a & \cdots & a & a & \cdots & x
\end{pmatrix}^{(n-1)} \qquad (\text{第 } l \text{ 行})$$

$$=xd_{n-1}+(1-n)a\det\begin{pmatrix} a & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{pmatrix}^{(n-1)}$$

$$=xd_{n-1}-(n-1)a^2(x-a)^{n-2}$$

由 $d_1=x$ 及递推公式

$$d_n=xd_{n-1}-a^2(n-1)(x-a)^{n-2},$$

有

$$d_2=xd_1-a^2(x-a)^0,$$

$$d_3=xd_2-2a^2(x-a),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$d_{n-1}=xd_{n-2}-(n-2)a^2(x-a)^{n-3},$$

$$d_n=xd_{n-1}-(n-1)a^2(x-a)^{n-2}.$$

依次乘以 $x^{n-2},x^{n-3},\cdots,x,1$，再全体相加，有

$$d_n=x^n-a^2\sum_{k=1}^{n-1}kx^{n-1-k}(x-a)^{k-1}$$

$$=x^n-[x^2-2x(x-a)+(x-a)^2]\sum_{k=1}^{n-1}kx^{n-1-k}(x-a)^{k-1}$$

$$=[x+(n-1)a](x-a)^{n-1}.$$

解 6

$$d_n=\det\begin{pmatrix} (x-a)+a & 0+a & \cdots & 0+a \\ 0+a & (x-a)+a & \cdots & 0+a \\ \vdots & \vdots & & \vdots \\ 0+a & 0+a & \cdots & (x-a)+a \end{pmatrix}$$

用定理 2.2.1 的性质 5，便有

$$d_n=(x-a)^n+\sum_{j=1}^{n}\det\begin{pmatrix} x-a & 0 & \cdots & \overset{(j)}{a} & \cdots & 0 \\ 0 & x-a & \cdots & a & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & a & \cdots & x-a \end{pmatrix}$$

$$=(x-a)^n+na(x-a)^{n-1}=[x+(n-1)a](x-a)^{n-1}.$$

习题 2.4

试求下列行列式的值:

1. $\det\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 3 & 4 & \cdots & n+1 \\ 3 & 4 & 5 & \cdots & n+2 \\ \vdots & \vdots & \vdots & & \vdots \\ n & n+1 & n+2 & \cdots & 2n-1 \end{pmatrix}$,

2. $\det\begin{pmatrix} x & -1 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & \cdots & 0 & 0 \\ 0 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & -1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 \end{pmatrix}$

3. $\det\begin{pmatrix} \sin\theta_1 & \sin 2\theta_1 & \cdots & \sin n\theta_1 \\ \sin\theta_2 & \sin 2\theta_2 & \cdots & \sin n\theta_2 \\ \vdots & \vdots & & \vdots \\ \sin\theta_n & \sin 2\theta_n & \cdots & \sin n\theta_n \end{pmatrix}$,

4. $\det\begin{pmatrix} a & 0 & & 0 & 1 \\ 0 & a & \vdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & a & 0 \\ 1 & 0 & \cdots & 0 & a \end{pmatrix}$,

5. 设 $a_{11}\neq 0$,试证:当 $n\geqslant 2$ 时有

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

$$=a_{11}^{2-n}\det\begin{pmatrix}\det\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix} & \cdots & \det\begin{pmatrix}a_{11} & a_{1n}\\ a_{21} & a_{2n}\end{pmatrix}\\ \vdots & & \vdots\\ \det\begin{pmatrix}a_{11} & a_{12}\\ a_{n1} & a_{n2}\end{pmatrix} & \cdots & \det\begin{pmatrix}a_{11} & a_{1n}\\ a_{n1} & a_{nn}\end{pmatrix}\end{pmatrix},$$

6. $\det\begin{pmatrix}a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n\\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n\\ \vdots & \vdots & & \vdots\\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n\end{pmatrix},$

7. $\det\begin{pmatrix}1+x_1y_1 & \cdots & 1+x_ny_1\\ \vdots & & \vdots\\ 1+x_1^ny_n & \cdots & 1+x_n^ny_n\end{pmatrix},$

8. $\det\begin{pmatrix}a_1 & & & & & & & b_{2n}\\ & a_2 & & & & & b_{2n-1} & \\ & & \ddots & & & \iddots & & \\ & & & a_n & b_{n+1} & & & \\ & & & b_n & a_{n+1} & & & \\ & & \iddots & & & \ddots & & \\ & b_2 & & & & & a_{2n-1} & \\ b_1 & & & & & & & a_{2n}\end{pmatrix},$

9. $\det\begin{pmatrix}a_1 & b & \cdots & b & b\\ c & a_2 & \cdots & b & b\\ \vdots & \vdots & & \vdots & \vdots\\ c & c & \cdots & a_{n-1} & b\\ c & c & \cdots & c & a_n\end{pmatrix},$ 10. $\det\begin{pmatrix}a_1 & b & \cdots & b & 1\\ c & a_2 & \cdots & b & 1\\ \vdots & \vdots & & \vdots & \vdots\\ c & c & \cdots & a_{n-1} & 1\\ 1 & 1 & \cdots & 1 & 0\end{pmatrix},$

11. $\det\begin{pmatrix}b_1 & a_2 & \cdots & a_n\\ a_1 & b_2 & \cdots & a_n\\ \vdots & \vdots & & \vdots\\ a_1 & a_2 & \cdots & b_n\end{pmatrix},$ 12. $\det\begin{pmatrix}\frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n}\\ \vdots & & \vdots\\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_n}\end{pmatrix},$

13. $\det\begin{pmatrix} \binom{m}{0} & \binom{m}{1} & \cdots & \binom{m}{n-1} \\ \vdots & \vdots & & \vdots \\ \binom{m+n-1}{0} & \binom{m+n-1}{1} & \cdots & \binom{m+n-1}{n-1} \end{pmatrix}, n-1\leqslant m,$

14. $\det\begin{pmatrix} \binom{m_1}{0} & \binom{m_1}{1} & \cdots & \binom{m_1}{n-1} \\ \vdots & \vdots & & \vdots \\ \binom{m_n}{0} & \binom{m_n}{1} & \cdots & \binom{m_n}{n-1} \end{pmatrix}, n-1\leqslant m_1<\cdots<m_n,$

15. $\sum\limits_{\binom{12\cdots n}{i_1 i_2\cdots i_n}} \det\begin{pmatrix} a_{1i_1} & \cdots & a_{1i_n} \\ \vdots & & \vdots \\ a_{ni_1} & \cdots & a_{ni_n} \end{pmatrix}$

16. 试证:

$$\det\begin{pmatrix} \binom{m+1}{m} & \binom{m+1}{m+1} & 0 & \cdots & 0 & 0 \\ \binom{m+2}{m} & \binom{m+2}{m+1} & \binom{m+2}{m+2} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \binom{m+n-2}{m} & \binom{m+n-2}{m+1} & \binom{m+n-2}{m+2} & \cdots & \binom{m+n-2}{m+n-2} & 0 \\ \binom{m+n-1}{m} & \binom{m+n-1}{m+1} & \binom{m+n-1}{m+2} & \cdots & \binom{m+n-1}{m+n-2} & \binom{m+n-1}{m+n-1} \\ \binom{m+n}{m} & \binom{m+n}{m+1} & \binom{m+n}{m+2} & \cdots & \binom{m+n}{m+n-2} & \binom{m+n}{m+n-1} \end{pmatrix}$$

$=\binom{m+n}{m},$

17. $\det\begin{pmatrix} x & 1 & & & & \\ -n & x-2 & 2 & & & \\ & -(n-1) & (x-4) & \ddots & & \\ & & \ddots & \ddots & & \\ & & & \ddots & x-2(n-1) & n \\ & & & & -1 & x-2n \end{pmatrix},$

18. $\det\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ -a_1 & a_2 & b_{23} & \cdots & b_{2n} \\ -a_1 & -a_2 & a_3 & \cdots & b_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ -a_1 & -a_2 & -a_3 & \cdots & b_{n-1,n} \\ -a_1 & -a_2 & -a_3 & \cdots & a_n \end{pmatrix}$,

19. $\det\begin{pmatrix} 1 & a_1 & \cdots & a_1^k & a_1^{k+2} & \cdots & a_1^n \\ 1 & a_2 & \cdots & a_2^k & a_2^{k+2} & \cdots & a_2^n \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & a_n & \cdots & a_n^k & a_n^{k+2} & \cdots & a_n^n \end{pmatrix}, 0\leqslant k\leqslant n-2$,

20. 设$a_{ij}, 1\leqslant i,j\leqslant n$ 都是实数，且 $a_{ii}>0, 1\leqslant i\leqslant n; a_{ij}<0, 1\leqslant i,j\leqslant n$, $i\neq j$，又$\sum_{j=1}^{n} a_{ij}>0, 1\leqslant i\leqslant n$. 试证 **Minkowski 定理**:

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}>0.$$

21. **Lovy-Desplanques 定理**: 设 $\det(a_{ij})=0$，则条件

$$2|a_{ii}|>\sum_{j=1}^{n}|a_{ij}|, \quad 1\leqslant i\leqslant n$$

不全成立，即 n 阶方阵 $A=(a_{ij})$ 不是**对角线占优方阵**.

22. 试证:

$$\det\begin{pmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{pmatrix}$$

为 λ 的 n 次多项式

$$\lambda^n-a_1\lambda^{n-1}+a_2\lambda^{n-2}-\cdots+(-1)^{n-1}a_{n-1}\lambda+(-1)^n a_n,$$

其中 a_j 为 $\det(a_{ij})$ 的所有 j 阶主子式之和，$j=1,2,\cdots,n$.

23. 试用归纳法证明 **Burnside 定理**: 斜对称方阵

$$A=(a_{ij}), \quad a_{ij}=-a_{ji}, \quad 1\leqslant i,j\leqslant n$$

的行列式有：当 n 为奇数，则 $\det A=0$；当 n 为偶数，则

$$\det A=\Bigg(\sum_{\substack{\binom{12\cdots n}{i_1 i_2\cdots i_n}\\ i_1<i_2,\cdots,i_{n-1}<i_n}} \delta^{12\cdots n}_{i_1 i_2\cdots i_n} a_{i_1 i_2} a_{i_3 i_4}\cdots a_{i_{n-1} i_n}\Bigg)^2.$$

§ 2.5 Cramer 法则

给定 mn 个数 a_{ij}, $1\leqslant i\leqslant n, 1\leqslant j\leqslant m$ 及 n 个数 $b_1,\cdots,b_n$. 记 $x_1,\cdots,x_m$ 为 m 个独立未知数,则 n 个方程

$$\begin{cases} a_{11}x_1+\cdots+a_{1m}x_m=b_1, \\ \cdots\cdots \\ a_{n1}x_1+\cdots+a_{nm}x_m=b_n. \end{cases}$$

构成**线性方程组**. m 个数 $x_1^{(0)},\cdots,x_m^{(0)}$ 称为它的**解**,如果有

$$\sum_{j=1}^{m} a_{ij}x_j^{(0)}=b_i,\ i=1,2,\cdots,n.$$

线性方程组的理论就是求解理论，它要解决下列问题：1. 什么时候解才存在？2. 如果解存在，用什么办法求出所有解？3. 什么时候解唯一存在？4. 如果解不存在，能不能求出某种意义的最小二乘解. 关于一般线性方程组的求解理论，放在第五章. 它回答了前三个问题,在第十二章回答第四个问题. 在这一节,我们讨论一种最简单的情形.

定理 2.5.1 (Cramer 法则) 设 n 阶行列式

$$\Delta=\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\neq 0,$$

则线性方程组

$$\sum_{i=1}^{n} a_{ij}x_j=b_i,\ 1\leqslant i\leqslant n$$

有一组唯一的解

$$x_j^{(0)}=\frac{\Delta_j}{\Delta},\quad 1\leqslant j\leqslant n$$

其中

$$\Delta_j=\sum_{k=1}^{n}b_kA_{kj}=\det\begin{pmatrix}a_{11}&\cdots&a_{1,j-1}&b_1&a_{1,j+1}&\cdots&a_{1n}\\ \vdots&&\vdots&\vdots&\vdots&&\vdots\\ a_{n1}&\cdots&a_{n,j-1}&b_n&a_{n,j+1}&\cdots&a_{nn}\end{pmatrix}.$$

证 今用 $x_j=\frac{\Delta_j}{\Delta}$, $1\leqslant j\leqslant n$ 代入方程组，有

$$\begin{aligned}\sum_{j=1}^{n}a_{ij}\frac{\Delta_j}{\Delta}&=\Delta^{-1}\sum_{j=1}^{n}a_{ij}\sum_{k=1}^{n}b_kA_{kj}\\&=\Delta^{-1}\sum_{k=1}^{n}b_k\left(\sum_{j=1}^{n}a_{ij}A_{kj}\right)\\&=\Delta^{-1}\sum_{k=1}^{n}b_k(\delta_{ik}\Delta)\\&=b_i,\quad i=1,2,\cdots n.\end{aligned}$$

所以 $x_j^{(0)}=\frac{\Delta_j}{\Delta}$, $1\leqslant j\leqslant n$ 为一组解. 下面再证唯一性，即任取一组解 $y_1^{(0)},\cdots,y_n^{(0)}$，我们要证 $y_j^{(0)}=x_j^{(0)}=\frac{\Delta_j}{\Delta}$, $1\leqslant j\leqslant n$. 事实上，由于 $y_1^{(0)},\cdots,y_n^{(0)}$ 为解，于是有

$$\sum_{j=1}^{n}a_{ij}y_j^{(0)}=b_i,\ i=1,2,\cdots,n.$$

因此由 Laplace 展开定理，有

$$\begin{aligned}x_j^{(0)}=\Delta^{-1}\Delta_j&=\Delta^{-1}\sum_{i=1}^{n}b_iA_{ij}\\&=\Delta^{-1}\sum_{i=1}^{n}\left(\sum_{k=1}^{n}a_{ik}y_k^{(0)}A_{ij}\right)\\&=\Delta^{-1}\sum_{k=1}^{n}y_k^{(0)}\sum_{i=1}^{n}a_{ik}A_{ij}\end{aligned}$$

$$= \Delta^{-1} \sum_{k=1}^{n} y_k^{(0)} \delta_{kj} \Delta$$

$$= y_j^{(0)}, \quad j = 1, 2, \cdots, n.$$

这证明了唯一性．证完．

定理 2.5.2 线性方程组

$$\sum_{j=1}^{n} a_{ij} x_j = 0, \quad i = 1, 2, \cdots, n$$

有非零解（即有一组解 $x_1^{(0)}$，$x_2^{(0)}$，…，$x_n^{(0)}$，它们不全为零），则 $\det(a_{ij}) = 0$．

证 用反证法，如果 $\Delta = \det(a_{ij}) \neq 0$，由定理 2.5.1，则有唯一解 $x_1^{(0)} = 0, \cdots, x_n^{(0)} = 0$．这和非零解的假设矛盾．证完．

习题 2.5

1. 试求两条平面曲线

$5x^2 - 8xy + 5y^2 + 26x - 28y + 32 = 0$，$2x^2 - 5xy + 2y^2 + 14x - 13y + 20 = 0$

的所有交点．

2. 求下列线性方程组的解，其中 $a_1, \cdots, a_n$ 为 n 个不同的非零数，

(i) $\sum_{i=1}^{n} a_i^{k-1} x_i = b_k, \quad k = 1, 2, \cdots, n;$

(ii) $\sum_{k=1}^{n} a_i^{k-1} x_k = b_i, \quad i = 1, 2, \cdots, n.$

第三章　多元多项式理论

§3.1　多元多项式和对称多项式

给定 n 个独立未知数 $x_1, x_2, \cdots, x_n$. 给定非负整数 $k_1, k_2, \cdots, k_n$ 及数 $a_{k_1k_2\cdots k_n}$, 则

$$a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$$

称为单项式, $a_{k_1k_2\cdots k_n}$ 称为系数, 而

$$k=k_1+k_2+\cdots+k_n$$

称为这个单项式的次数.

定义　有限个单项式之和称为 **n 元多项式**. n 元多项式 f 的**次数**为这有限个非零单项式的次数之最大值, 记作 $\deg(f)$.

给定 n 元多项式 $f(x_1, x_2, \cdots, x_n)$, 我们可以用很多种方式将它排成单项式之和. 最常用的排法, 称为**字典排列法**. 定义如下: 给定两个单项式

$$x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}, \quad x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}.$$

如果有

$$k_1=j_1, \quad k_2=j_2, \quad \cdots, \quad k_s=j_s, \quad k_{s+1}<j_{s+1},$$

则我们说 $x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$ 在 $x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 之前. 这时根本不用考虑 $j_{s+2}, \cdots, j_n$ 与 $k_{s+2}, \cdots, k_n$ 之间的关系.

字典排列法的缺陷在于次数最高的单项式不一定排在前面. 为了克服这个困难, 我们用两种方法结合起来排.

给定 n 元多项式 $f(x_1, \cdots, x_n)$, 将其中次数相同的单项式放在一起, 构成一个新的多元多项式, 它的每个非零单项式的次数等于

这个多元多项式的次数，这种多元多项式称为**齐次多项式**．设 $f_s(x_1,\cdots,x_n)$ 为 s 次齐次多项式，于是有：对一切参数 t，则

$$f_s(tx_1,tx_2,\cdots,tx_n)=t^s f_s(x_1,\cdots,x_n),\quad s=0,1,2,\cdots.$$

反之，立即可证：适合上式之多元多项式必为 s 次齐次多项式．且对任一多元多项式 f，则有

$$f=f_m+f_{m-1}+\cdots+f_1+f_0$$

其中 $f_m\neq 0$, $m=\deg(f)$．再对每个 s 次齐次多项式 f_s 用字典排列法排成单项式之和，合并在一起，便将多元多项式 f 排成了单项式之和．

显然，零次齐次多项式为一个非零常数．又零本身称为**零多项式**．

和一个未知数 x 的情形相同，我们可以很自然地引进两个 n 元多项式的和、减和乘这三种运算，所有运算性质也和一元情形完全相同．

引理 3.1.1 将非零 n 元多项式 $f(x_1,\cdots,x_n)$ 和 $g(x_1,\cdots,x_n)$ 分别按字典排列法排好．首项分别记作

$$a_0x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n},\quad b_0x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$$

将乘积 fg 也按字典排列法排好，则其首项为

$$a_0b_0x_1^{j_1+k_1}x_2^{j_2+k_2}\cdots x_n^{j_n+k_n}.$$

证 注意到 fg 为 f 中所有单项式和 g 的所有单项式相乘，然后归并同类项，再重新按字典排列法排好．今 f 中任一单项式 $ax_1^{p_1}x_2^{p_2}\cdots x_n^{p_n}$ 和 g 中任一单项式 $bx_1^{q_1}x_2^{q_2}\cdots x_n^{q_n}$ 相乘，为

$$abx_1^{p_1+q_1}x_2^{p_2+q_2}\cdots x_n^{p_n+q_n}.$$

由于 $a_0x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$ 为 f 之首项，所以有

$$j_1=p_1,\cdots,j_s=p_s,\ j_{s+1}>p_{s+1}\quad 或\quad j_1=p_1,\cdots,j_n=p_n.$$

又 $b_0x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 为 g 之首项，所以有

$$k_1=q_1,\cdots,k_t=q_t,k_{t+1}>q_{t+1}\quad 或\quad k_1=q_1,\cdots,k_n=q_n.$$

显然，为讨论方便起见，无妨设 $n\geqslant s\geqslant t\geqslant 0$，于是当 $ax_1^{p_1}x_2^{p_2}\cdots x_n^{p_n}$ 和 $bx_1^{q_1}x_2^{q_2}\cdots x_n^{q_n}$ 中至少有一个不是首项时，必有

$$j_1+k_1=p_1+q_1,\cdots,j_t+k_t=p_t+q_t,j_{t+1}+k_{t+1}>p_{t+1}+q_{t+1}.$$

这证明了 $a_0b_0x_1^{j_1+k_1}x_2^{j_2+k_2}\cdots x_n^{j_n+k_n}$ 为首项. 证完.

引理 3.1.2 任给 n 元多项式 $f(x_1,\cdots,x_n)$ 和 $g(x_1,x_2,\cdots,x_n)$，则有

$$\deg(f+g)\leqslant\max(\deg(f),\deg(g)),$$

$$\deg(fg)=\deg(f)+\deg(g).$$

证 前一不等式显然，后一不等式证明如下：先将多元多项式 f 及 g 分别写成齐次多项式之和

$$f=f_p+f_{p-1}+\cdots+f_0,\quad g=g_q+g_{q-1}+\cdots+g_0.$$

于是

$$fg=\sum_{j=0}^{p}f_j\sum_{k=0}^{q}g_k=\sum_{j=0}^{p}\sum_{k=0}^{q}f_jg_k=\sum_{l=0}^{p+q}\Big(\sum_{j+k=l}f_jg_k\Big)=\sum_{l=0}^{p+q}h_l,$$

其中

$$h_l=\sum_{j+k=l}f_jg_k.$$

显然 j 次齐次多项式 f_j 和 k 次齐次多项式 g_k 之乘积 f_jg_k 为 $j+k$ 次齐次多项式，所以 h_l 如果不等于零，则必为 l 次齐次多项式. 今 $\deg f=p$，所以 $f_p\neq 0$，又 $\deg g=q$，所以 $g_q\neq 0$. 因此 $f_pg_q\neq 0$，于是 $h_{p+q}=f_pg_q\neq 0$，这证明了 $\deg(fg)=\deg(f)+\deg(g)$. 证完.

最重要的齐次多项式为下面两类.

(1) Newton 等幂和：取定独立未知数 $x_1,x_2,\cdots,x_n$ 及非负整数 k，则

$$s_k=x_1^k+x_2^k+\cdots+x_n^k,\quad k=0,1,2,\cdots$$

称为 **Newton 等幂和**. 于是 $s_0=n$，又 $\deg s_k=k,k=0,1,2,\cdots$. 再 s_k 之首项为 x_1^k

(2) 初等对称多项式：取定独立未知数 $x_1, x_2, \cdots, x_n$，

$$\sigma_1 = x_1 + x_2 + \cdots + x_n = \sum_{j=1}^{n} x_j,$$

$$\sigma_2 = x_1x_2 + \cdots + x_1x_n + \cdots + x_{n-1}x_n = \sum_{1 \leqslant j < k \leqslant n} x_j x_k,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sigma_{n-1} = x_1x_2\cdots x_{n-1} + x_1x_2\cdots x_{n-2}x_n + \cdots + x_2x_3\cdots x_n$$

$$= \sum_{1 \leqslant j_1 < \cdots < j_{n-1} \leqslant n} x_{j_1}x_{j_2}\cdots x_{j_{n-1}}.$$

$$\sigma_n = x_1x_2\cdots x_n.$$

它们分别为 $1,2,\cdots,n$ 次齐次多项式，称为**初等对称多项式**.

易证：对独立未知数 $x, x_1, \cdots, x_n$，则有 **Vieta 定理**：

$$(x-x_1)(x-x_2)\cdots(x-x_n) = x^n - \sigma_1 x^{n-1} + \sigma_2 x^{n-2} - \cdots + (-1)^{n-1}\sigma_{n-1}x + (-1)^n\sigma_n.$$

且对 $1,2,\cdots,n$ 的任一排列 $i_1i_2\cdots i_n$，则由

$$(x-x_1)(x-x_2)\cdots(x-x_n) = (x-x_{i_1})(x-x_{i_2})\cdots(x-x_{i_n}),$$

所以证明了

$$\sigma_j(x_{i_1}, x_{i_2}, \cdots, x_{i_n}) = \sigma_j(x_1, x_2, \cdots, x_n), \quad j = 1, 2, \cdots, n,$$

将这个性质提炼为

定义 n 元多项式 $f(x_1, x_2, \cdots, x_n)$ 称为**对称多项式**，如果对 $1,2,\cdots,n$ 的所有排列 $i_1i_2\cdots i_n$，都有

$$f(x_{i_1}, x_{i_2}, \cdots, x_{i_n}) = f(x_1, x_2, \cdots, x_n)$$

由定义可知初等对称多项式 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 都是对称多项式，又 Newton 等幂和 $s_0, s_1, s_2, \cdots$ 也都是对称多项式.

引理 3.1.3 将对称多项式 $f(x_1, \cdots, x_n)$ 分解为齐次多项式之和 $f = \sum_{j=0}^{m} f_j$，则每个齐次多项式 f_j 仍为对称多项式.

证 对 $1,2,\cdots,n$ 的任一排列 $i_1i_2\cdots i_n$，则

$$\sum_{j=0}^{m} f_j(x_{i_1}, x_{i_2}, \cdots, x_{i_n}) = \sum_{j=0}^{m} f_j(x_1, x_2, \cdots, x_n).$$

比较同次齐次多项式，便证明了 $f_j(x_{i_1}, x_{i_2}, \cdots, x_{i_n}) = f_j(x_1, x_2, \cdots, x_n)$. 这证明了 f_i 是对称多项式. 证完.

显然有

引理 3.1.4 初等对称多项式 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的多项式

$$F(\sigma_1, \sigma_2, \cdots, \sigma_n) = f(x_1, x_2, \cdots, x_n)$$

为 $x_1, x_2, \cdots, x_n$ 的对称多项式.

反之，有

定理 3.1.1 (对称多项式基本定理) 对称多项式可唯一地表为初等对称多项式的多项式.

证 将对称多项式 $f(x_1, x_2, \cdots, x_n)$ 按照字典排列法排成单项式之和，设首项为

$$a_{j_1 j_2 \cdots j_n} x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}.$$

先来证明 $j_1 \geqslant j_2 \geqslant \cdots \geqslant j_n$. 设若不然，则存在指标 $j_l < j_{l+1}$. 因为 $f(x_1, x_2, \cdots, x_n)$ 为对称多项式，所以

$$f(x_1, \cdots, x_{l-1}, x_{l+1}, x_l, x_{l+2}, \cdots, x_n) = f(x_1, \cdots, x_n).$$

因此 $f(x_1, x_2, \cdots, x_n)$ 作为单项式之和，出现一项

$$a_{j_1 j_2 \cdots j_n} x_1^{j_1} \cdots x_{l-1}^{j_{l-1}} x_l^{j_{l+1}} x_{l+1}^{j_l} x_{l+2}^{j_{l+2}} \cdots x_n^{j_n}.$$

然而按照字典排列法，它不是首项. 所以 $j_l \geqslant j_{l+1}$. 这和假设 $j_l < j_{l+1}$ 矛盾. 这证明了 $j_1 \geqslant j_2 \geqslant \cdots \geqslant j_n$.

作对称多项式

$$\varphi_1(x_1, x_2, \cdots, x_n) = a_{j_1 j_2 \cdots j_n} \sigma_1^{j_1 - j_2} \sigma_2^{j_2 - j_3} \cdots \sigma_{n-1}^{j_{n-1} - j_n} \sigma_n^{j_n},$$

由引理 3.1.1，多项式 φ_1 的首项为

$$a_{j_1 j_2 \cdots j_n} x_1^{j_1 - j_2} (x_1 x_2)^{j_2 - j_3} \cdots (x_1 x_2 \cdots x_{n-1})^{j_{n-1} - j_n} (x_1 x_2 \cdots x_n)^{j_n}$$
$$= a_{j_1 j_2 \cdots j_n} x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}.$$

所以多项式 f 和 φ_1 有相同的首项. 因此对称多项式

$$f_1(x_1,x_2,\cdots,x_n)=f(x_1,x_2,\cdots,x_n)-\varphi_1(x_1,x_2,\cdots,x_n)$$

的项，按字典排列法排好后，都在$a_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$之后. 所以首项$a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$有$k_1\geqslant k_2\geqslant\cdots\geqslant k_n$，且

$$j_1=k_1,\cdots,j_r=k_r,\quad j_{r+1}>k_{r+1}\geqslant k_{r+2}\geqslant\cdots\geqslant k_n,$$

这里$0\leqslant r<n$. 于是可构作$a_{k_1k_2\cdots k_n}\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\cdots\sigma_{n-1}^{k_{n-1}-k_n}\sigma_n^{k_n}$，这样依次讨论下去. 由于适合$j_1\geqslant p_1\geqslant\cdots\geqslant p_n$的$n$数组$(p_1,p_2,\cdots,p_n)$只有有限组，所以上述作法必然在有限步后结束. 这证明了

$$f=\varphi_1+\varphi_2+\cdots+\varphi_s,$$

其中φ_j为$\sigma_1,\sigma_2,\cdots,\sigma_n$的单项式. 这证明了对称多项式$f$必为$\sigma_1,\sigma_2,\cdots,\sigma_n$的多项式.

下面来证明对称多项式$f(x_1,x_2,\cdots,x_n)$表成初等对称多项式$\sigma_1,\sigma_2,\cdots,\sigma_n$的多项式$F(\sigma_1,\sigma_2,\cdots,\sigma_n)$时，表法唯一，即证明两个初等对称多项式$\sigma_1,\sigma_2,\cdots,\sigma_n$的多项式$g_1(\sigma_1,\cdots,\sigma_n)$和$g_2(\sigma_1,\cdots,\sigma_n)$若作为$x_1,x_2,\cdots,x_n$的多项式是相等的，则作为$\sigma_1,\sigma_2,\cdots,\sigma_n$的多项式也是相等的. 换句话说，只要证明$\sigma_1,\sigma_2,\cdots,\sigma_n$的多项式$g(\sigma_1,\sigma_2,\cdots,\sigma_n)$作为$x_1,x_2,\cdots,x_n$的多项式等于零，那末作为$\sigma_1,\sigma_2,\cdots,\sigma_n$的多项式也等于零. 用反证法来证明. 如果$g(\sigma_1,\sigma_2,\cdots,\sigma_n)$不是零多项式，作为$x_1,x_2,\cdots,x_n$的多项式它为零多项式. 显然$g$不可能是单项式，因此$g$中至少有两个不同单项式. 任取两个不同单项式$a_{j_1j_2\cdots j_n}\sigma_1^{j_1}\sigma_2^{j_2}\cdots\sigma_n^{j_n}$，$a_{k_1k_2\cdots k_n}\sigma_1^{k_1}\sigma_2^{k_2}\cdots\sigma_n^{k_n}$. 它们作为$x_1,x_2,\cdots,x_n$的多项式，首项各为

$$a_{j_1j_2\cdots j_n}x_1^{j_1}(x_1x_2)^{j_2}\cdots(x_1x_2\cdots x_n)^{j_n},$$

$$a_{k_1k_2\cdots k_n}x_1^{k_1}(x_1x_2)^{k_2}\cdots(x_1x_2\cdots x_n)^{k_n}$$

此即

$$a_{j_1j_2\cdots j_n}x_1^{j_1+j_2+\cdots+j_n}x_2^{j_2+\cdots+j_n}\cdots x_n^{j_n},\ a_{k_1k_2\cdots k_n}x_1^{k_1+k_2+\cdots+k_n}x_2^{k_2+\cdots+k_n}\cdots x_n^{k_n}.$$

这是同类项当且仅当

$$j_1+j_2+\cdots+j_n=k_1+k_2+\cdots+k_n,$$

$$j_2+\cdots+j_n=k_2+\cdots+k_n, \cdots, j_n=k_n.$$

所以 $j_1=k_1, j_2=k_2, \cdots, j_n=k_n$. 这证明了 $g(\sigma_1, \cdots, \sigma_n)$的每个非零单项式作为 $x_1, x_2, \cdots, x_n$ 之多项式，首项都不相同．而这些首项再按字典排列法排出之首项，必然为 $g(\sigma_1, \cdots, \sigma_n)$作为 $x_1, x_2, \cdots, x_n$ 的多项式按字典排列法排出之首项．所以 $g(\sigma_1, \sigma_2, \cdots, \sigma_n)$作为 $x_1, x_2, \cdots, x_n$ 的多项式，有非零首项，即不是零多项式．这和假设矛盾．即证明了 $g(\sigma_1, \sigma_2, \cdots, \sigma_n)$作为 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的多项式为零多项式．证完．

由这个定理的证明，同时给出了如何将对称多项式表为初等对称多项式的多项式的计算方案．

（1）将已给的对称多项式写成齐次多项式之和．然后对每个齐次多项式进行计算．

（2）设 $f(x_1, x_2, \cdots, x_n)$为 m 次齐次对称多项式，找出所有单项式 $a_{i_1 i_2 \cdots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$，使得 $i_1 \geqslant i_2 \geqslant \cdots \geqslant i_n$.

（3）令

$$f(x_1, x_2, \cdots, x_n)=\sum_{i_1 \geqslant \cdots \geqslant i_n} A_{i_1 i_2 \cdots i_n} \sigma_1^{i_1-i_2} \sigma_2^{i_2-i_3} \cdots \sigma_{n-1}^{i_{n-1}-i_n} \sigma_n^{i_n}.$$

再用 $x_1, \cdots, x_n$ 的特殊值代入，以便求系数 $A_{i_1 i_2 \cdots i_n}$. 代入的特殊值例如可用 $x_1=1, \cdots, x_r=1, x_{r+1}=0, \cdots, x_n=0$ 等等．

习题 3.1

1．将对称多项式 $f(x_1, x_2, \cdots, x_n)$及 $\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}$分别表为初等对称多项式的多项式时，它们之间有什么关系？

2．设多元多项式 $f(x_1, x_2, \cdots, x_n)$ 在未知数 $x_1, x_2, \cdots, x_n$ 的所有偶排列下不变，试证：

$$f(x_1, x_2, \cdots, x_n)=g(x_1, x_2, \cdots, x_n)+h(x_1, x_2, \cdots, x_n) V(x_1, x_2, \cdots, x_n),$$

其中 g 和 h 为对称多项式，$V(x_1, \cdots, x_n)$为 Vandermonde 行列式．

3. 未知数 $x_1, x_2, \cdots, x_n$ 的轮回排列为

$$x_1x_2\cdots x_n,\quad x_2x_3\cdots x_nx_1,\quad x_3x_4\cdots x_nx_1x_2, \cdots, x_nx_1x_2\cdots x_{n-1}.$$

设多项式 $f(x_1, x_2, \cdots, x_n)$ 在未知数 $x_1, x_2, \cdots, x_n$ 的轮回排列下不变，试证：

$$f(x_1, \cdots, x_n) = \sum a_{i_0 i_1 \cdots i_{n-1}} g_0^{i_0} g_1^{i_1} \cdots g_n^{i_{n-1}},$$

其中和取遍所有适合条件

$$n \mid (i_1 + 2i_2 + \cdots + (n-1)i_{n-1}),\quad i_0 + i_1 + \cdots + i_{n-1} \leqslant \deg f,$$

的非负整数组 $i_0, i_1, \cdots, i_{n-1}$，又

$$g_j = x_1\varepsilon^j + x_2\varepsilon^{2j} + \cdots + x_n\varepsilon^{nj},\quad j = 0, 1, 2, \cdots, n-1,$$

其中 $\varepsilon = \cos\frac{2\pi}{n} + \sqrt{-1}\sin\frac{2\pi}{n}$ 为 $x^n = 1$ 之本原单位根.

4. 试证：n^2 个独立未知数 $x_{ij}, 1 \leqslant i, j \leqslant n$ 的多项式

$$\det\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}$$

不可约.

5. 试求所有适合条件 $f(1, 0) = 1$,

$$f(x+y, z) + f(y+z, x) + f(z+x, y) = 0$$

的 n 次齐次多项式 $f(x, y)$.

§3.2 结式和判别式

给定两个一元多项式

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n,\quad g(x) = b_0x^m + b_1x^{m-1} + \cdots + b_m$$

假设它们次数都大于零，且 $a_0 \neq 0, b_0 \neq 0$. 由 §1.4 可知 f 和 g 有公根当且仅当它们不互素. 由定理 1.2.2 可知 f 和 g 有公根当且仅当存在非零多项式 u, v 使得

$$uf = vg.$$

且 $0 \leqslant \deg(u) < m, 0 \leqslant \deg(v) < n$. 将 u, v 写成

$$u = x_{n+1}x^{m-1} + x_{n+2}x^{m-2} + \cdots + x_{n+m},$$

$$v = x_1x^{n-1} + x_2x^{n-2} + \cdots + x_n,$$

其中 $x_1, \cdots, x_n$ 不全为零, $x_{n+1}, \cdots, x_{n+m}$ 也不全为零.

引理 3.2.1 符号同上. 多项式 f 和 g 有公根当且仅当 $n+m$ 个未知数 $x_1, x_2, \cdots, x_{n+m}$ 的线性方程组

$b_l x_1 + \cdots + b_{l-n+1} x_n - a_l x_{n+1} - \cdots - a_{l-m+1} x_{n+m} = 0, l = 0, 1, 2, \cdots, n+m-1$ 有非零解, 其中当 $j \neq 0, 1, \cdots, n$ 时 $a_j = 0$, 当 $j \neq 0, 1, \cdots, m$ 时 $b_j = 0$.

证 上面讨论告诉我们, f 和 g 有公根当且仅当存在 $x_1^{(0)}, \cdots, x_n^{(0)}$ 不全为零, 又存在 $x_{n+1}^{(0)}, \cdots, x_{n+m}^{(0)}$ 不全为零, 使得

$$
\begin{aligned}
uf - vg &= \left(\sum_{j=0}^{n} a_j x^{n-j}\right)\left(\sum_{k=0}^{m-1} x_{n+k+1}^{(0)} \ x^{m-k-1}\right) \\
&\quad - \left(\sum_{p=0}^{m} b_p x^{m-p}\right)\left(\sum_{q=0}^{n-1} x_{q+1}^{(0)} x^{n-q-1}\right) \\
&= \sum_{l=0}^{n+m-1} \left(\sum_{j+k=l} a_j x_{n+k+1}^{(0)} - \sum_{p+q=l} b_p x_{q+1}^{(0)}\right) x^{n+m-l-1} \\
&= 0,
\end{aligned}
$$

此即 $x_1^{(0)}, \cdots, x_{n+m}^{(0)}$ 为线性方程组

$$
\sum_{j+k=l} a_j x_{n+k+1} = \sum_{p+q=l} b_p x_{q+1}, \quad l = 0, 1, \cdots, n+m-1
$$

的解. 所以为了证明引理, 只要证明上述线性方程组若有非零解 $x_1^{(0)}, \cdots, x_{n+m}^{(0)}$, 则 $x_1^{(0)}, \cdots, x_n^{(0)}$ 不全为零, 且 $x_{n+1}^{(0)}, \cdots, x_{n+m}^{(0)}$ 不全为零. 为此用反证法. 若 $x_1^{(0)} = \cdots = x_n^{(0)} = 0$, 则有

$$
a_l x_{n+1}^{(0)} + \cdots + a_{l-m+1} x_{n+m}^{(0)} = 0, \quad l = 0, 1, \cdots, n+m-1
$$

由 $a_0 \neq 0$, 便证明了 $x_{n+1}^{(0)} = \cdots = x_{n+m}^{(0)} = 0$. 这导出矛盾. 若 $x_{n+1}^{(0)} = \cdots = x_{n+m}^{(0)} = 0$, 则有

$$
b_l x_1^{(0)} + \cdots + b_{l-n+1} x_n^{(0)} = 0, \quad l = 0, 1, \cdots, n+m-1.
$$

由 $b_0 \neq 0$ 便证明了 $x_1^{(0)} = \cdots = x_n^{(0)} = 0$, 这也导出矛盾. 证完.

定义 非零多项式

$$
f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n, \quad g(x) = b_0 x^m + b_1 x^{m-1} + \cdots + b_m
$$

决定的 $n+m$ 阶行列式

$$R(f,g)=\det\begin{pmatrix} a_0 & a_1 & \cdots & a_n & & & \\ & a_0 & a_1 & \cdots & a_n & & \\ & & \ddots & \ddots & & \ddots & \\ & & & a_0 & a_1 & \cdots & a_n \\ b_0 & \cdots & b_{m-1} & b_m & & & \\ & b_0 & \cdots & b_{m-1} & b_m & & \\ & & \ddots & & \ddots & \ddots & \\ & & & b_0 & \cdots & b_{m-1} & b_m \end{pmatrix}\begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m\text{行} \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n\text{行}\end{matrix}$$

称为 f 和 g 的**结式**.

引理 3.2.2 次数大于零的多项式 $f(x)$ 和 $g(x)$ 有公根当且仅当它们的结式 $R(f,g)=0$.

证 引理 3.2.1 给出的线性方程组的系数行列式等于 $(-1)^n R(f,g)$. 引理 3.2.1 证明了 f 和 g 有公根当且仅当上述线性方程组有非零解. 由 §2.5可知当且仅当系数行列式等于零, 即 $R(f,g)=0$. 证完.

定理 3.2.1 设次数大于零的多项式 $f(x)$ 和 $g(x)$ 分别有根 $x_1,x_2,\cdots,x_n$ 及 $y_1,y_2,\cdots,y_m$. 则 f 和 g 的结式

$$R(f,g)=a_0^m b_0^n\prod_{j=1}^{n}\prod_{k=1}^{m}(x_j-y_k)=a_0^m\prod_{j=1}^{n}g(x_j)$$

$$=(-1)^{nm}b_0^n\prod_{k=1}^{m}f(y_k),$$

其中 a_0 和 b_0 分别为 f 和 g 的首项系数

证 由结式的定义可知

$$R(\lambda f,\mu g)=\lambda^m\mu^n R(f,g),$$

对一切常数 λ,μ 成立. 因此设 $f=a_0x^n+\cdots+a_n$, $g=b_0x^m+\cdots+b_m$, 取 $\lambda=a_0^{-1}$, $\mu=b_0^{-1}$, 从而无妨设 $a_0=1$, $b_0=1$. 即有

$$f(x)=(x-x_1)(x-x_2)\cdots(x-x_n)=x^n+a_1x^{n-1}+\cdots+a_n$$

$$g(x)=(x-y_1)(x-y_2)\cdots(x-y_m)=x^m+b_1x^{m-1}+\cdots+b_m$$

记 $x_1,x_2,\cdots,x_n$ 之初等对称多项式为 $\sigma_1,\cdots,\sigma_n$，由 Vieta 定理，于是

$$a_j=(-1)^j\sigma_j,\quad j=1,2,\cdots,n.$$

又记 $y_1,y_2,\cdots,y_m$ 之初等对称多项式为 $\tau_1,\cdots,\tau_m$，同理

$$b_k=(-1)^k\tau_k,\quad k=1,2,\cdots,m.$$

由行列式的定义可知结式 $R(f,g)$ 为初等对称多项式 $\sigma_1,\sigma_2,\cdots,\sigma_n$ 的多项式，也是初等对称多项式 $\tau_1,\tau_2,\cdots,\tau_m$ 的多项式. 注意到当 $x_2,\cdots,x_n,y_1,\cdots,y_m$ 取定时，$R(f,g)$ 为 x_1 的多项式. 令 $x_1=y_k(1\leqslant k\leqslant m)$，则由于这时 f 和 g 有公根 $x_1=y_k$，由引理 3.2.2，x_1 的多项式 $R(f,g)$ 之值为零，即以 $y_1,\cdots,y_m$ 为根. 于是有 $(x_1-y_k)\mid R(f,g)$，这证明了

$$R(f,g)=h(x_2,\cdots,x_n;y_1,\cdots,y_m)\prod_{k=1}^{m}(x_1-y_k)$$

比较双方 x_1^m 之系数，便证明了 $h(x_2,\cdots,x_n;\ y_1,\cdots,y_m)$ 为 $x_2,\cdots,x_n;y_1,\cdots,y_m$ 之多项式. 对 $x_2,\cdots,x_n$ 同样讨论，最后证明了

$$R(f,g)=h_0\prod_{j=1}^{n}\prod_{k=1}^{m}(x_j-y_k),$$

其中 h_0 为常数.

为了确定常数 h_0，取 $g=x^m$，即取 $b_0=1,b_1=\cdots=b_m=0$，于是 $y_1,\cdots,y_n$ 都为零. 因此

$$R(f,g)=h_0\prod_{j=1}^{n}x_j^m=h_0(x_1x_2\cdots x_n)^m=h_0\sigma_n^m.$$

由直接计算行列式 $R(f,g)$ 便有

$$R(f,g)=(-1)^{nm}a_n^m=((-1)^na_n)^m=\sigma_n^m.$$

这证明了 $h_0=1$. 定理证完.

定义 给定次数大于零的多项式 $f(x)$，记作

$$f(x)=a_0(x-x_1)(x-x_2)\cdots(x-x_n),$$

则 $x_1,x_2,\cdots,x_n$ 的对称多项式

$$\Delta(f)=(-1)^{\frac{n(n-1)}{2}}a_0^{2n-1}\prod_{j\neq k}(x_j-x_k)=a_0^{2n-1}\prod_{1\leq j<k\leq n}(x_j-x_k)^2$$

称为 f 的**判别式**.

显然有

定理 3.2.2 次数大于零的多项式 $f(x)$ 有重根当且仅当 f 的判别式 $\Delta(f)=0$.

定理 3.2.3 对次数大于零的多项式 $f(x)$，有

$$\Delta(f)=(-1)^{\frac{n(n-1)}{2}}R(f,f').$$

证 令

$$f(x)=a_0(x-x_1)(x-x_2)\cdots(x-x_n),$$

$$f'(x)=a_0\sum_{l=1}^{n}(x-x_1)\cdots(x-x_{l-1})(x-x_{l+1})\cdots(x-x_n).$$

由定理 3.2.1，

$$R(f,f')=a_0^{n-1}\prod_{j=1}^{n}f'(x_j)=a_0^{2n-1}\prod_{j=1}^{n}\prod_{k\neq j}(x_j-x_k)$$

$$=(-1)^{\frac{n(n-1)}{2}}\Delta(f).$$

证完.

习题 3.2

1. 试证:

(i) $R(f,g_1g_2)=R(f,g_1)R(f,g_2)$,

(ii) $\Delta((x-a)f(x))=\Delta(f(x))(f(a))^2$,

(iii) $\Delta(f(g(x)))=\Delta((f))^{\deg(g)}\prod_{j=1}^{\deg(f)}\Delta(g(x)-x_j)$,

其中 $$f(x)=(x-x_1)(x-x_2)\cdots(x-x_n),$$

又 $g(x)$ 之首项为 1.

2. 已知 $\Delta(f(x))$，试求 $\Delta(f(x^m))$，$m=2$ $3,\cdots$.

第四章　矩阵的代数运算

§4.1　矩阵的代数运算

定义　nm 个元素排成的矩形

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

称为 $n\times m$ 矩阵. 两个矩阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & \cdots & b_{1q} \\ \vdots & & \vdots \\ b_{p1} & \cdots & b_{pq} \end{pmatrix}$$

称为**相等的**,如果 $p=n, q=m$,且 $b_{ij}=a_{ij}, 1\leqslant i\leqslant n, 1\leqslant j\leqslant m$. 给定 $n\times m$ 矩阵 $A=(a_{ij})$,则 $m\times n$ 矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1m} & \cdots & a_{nm} \end{pmatrix}$$

称为 A 的**转置矩阵**,它的第 i 行是 A 的第 i 列,它的第 j 列是 A 的第 j 行. 这个矩阵记作 A' 或 A^T.

下面在矩阵之间引进代数运算.

(1) 加法

任给两个 $n\times m$ 矩阵 A 和 B,则可定义它们的**和**

$$A+B=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}+\begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix}$$

$$=\begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1m}+b_{1m} \\ \vdots & & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{nm}+b_{nm} \end{pmatrix}.$$

由定义可知加法运算有性质:

(i) 加法交换律

设 A 和 B 都是 $n\times m$ 矩阵,则有

$$A+B=B+A$$

(ii) 加法结合律

设 A 和 B C 都是 $n\times m$ 矩阵,则有

$$(A+B)+C=A+(B+C).$$

(iii) 存在 $n\times m$ 矩阵

$$0=\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

称为**零矩阵**,使得对一切 $n\times m$ 矩阵 A 都有

$$A+0=0+A=A$$

(iv) 对每个 $n\times m$ 矩阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix},$$

则存在 $n\times m$ 矩阵

$$\begin{pmatrix} -a_{11} & \cdots & -a_{1m} \\ \vdots & & \vdots \\ -a_{n1} & \cdots & -a_{nm} \end{pmatrix}$$

称为 A 的**负矩阵**,记作 $-A$. 它有性质

$$(-A)+A=A+(-A)=0$$

因此可以引进减法

$$A-B=A+(-B)$$

从加法定义出发，容易证明，对任一 $n\times m$ 矩阵 A，若 $n\times m$ 矩阵 B 使 $A+B=B+A=A$，则 $B=0$. 又对任一 $n\times m$ 矩阵 A，若 $n\times m$ 矩阵 C 使 $A+C=C+A=0$，则 $C=-A$. 即零矩阵和每个矩阵的负矩阵都唯一存在.

(2) 纯量乘积

任给常数 λ 及 $n\times m$ 矩阵 A，则可定义它们的**纯量乘积** λA 如下：

$$\lambda A=\lambda\begin{pmatrix}a_{11} & \cdots & a_{1m}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nm}\end{pmatrix}=\begin{pmatrix}\lambda a_{11} & \cdots & \lambda a_{1m}\\ \vdots & & \vdots\\ \lambda a_{n1} & \cdots & \lambda a_{nm}\end{pmatrix}.$$

由定义可知纯量乘积有性质

(i) $(\lambda+\mu)A=\lambda A+\mu A$,

(ii) $\lambda(A+B)=\lambda A+\lambda B$,

(iii) $(\lambda\mu)A=\lambda(\mu A)$,

(iv) $1\cdot A=A,(-1)\cdot A=-A$,

(v) $\lambda A=0$ 当且仅当 $\lambda=0$ 或者 $A=0$,

其中 λ,μ 都是数，A,B 都是 $n\times m$ 矩阵.

引进 $n\times m$ 矩阵 E_{ij}，使得它的第 i 行，第 j 列元素为1，其余位置的元素为 0，即

$$E_{ij}=\begin{pmatrix}0\cdots & 0 & 0 & 0\cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 0\cdots & 0 & 0 & 0\cdots & 0\\ 0\cdots & 0 & 1 & 0\cdots & 0\\ 0\cdots & 0 & 0 & 0\cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 0\cdots & 0 & 0 & 0\cdots & 0\end{pmatrix}\begin{matrix}\\ \\ \\ (i),\\ \\ \\ \\ \end{matrix}\quad i=1,2,\cdots,n,\ j=1,2,\cdots,m.$$
$$\qquad\qquad(j)$$

由加法及纯量乘积的定义可知

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}=\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}E_{ij},$$

$$B=\begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix}=\sum_{i=1}^{n}\sum_{j=1}^{m}b_{ij}E_{ij},$$

则有

$$A+B=\Sigma a_{ij}E_{ij}+\Sigma b_{ij}E_{ij}=\Sigma(a_{ij}+b_{ij})E_{ij},$$
$$\lambda A=\lambda\Sigma a_{ij}E_{ij}=\Sigma(\lambda a_{ij})E_{ij}.$$

(3) 乘法

任给 $n\times m$ 矩阵 A 和 $m\times p$ 矩阵 B，则 A 和 B 的乘积 AB 定义为

$$AB=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}\begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mp} \end{pmatrix}=\begin{pmatrix} \sum_{j=1}^{m}a_{1j}b_{j1} & \cdots & \sum_{j=1}^{m}a_{1j}b_{jp} \\ \vdots & & \vdots \\ \sum_{j=1}^{m}a_{nj}b_{j1} & \cdots & \sum_{j=1}^{m}a_{nj}b_{jp} \end{pmatrix}$$

所以乘积为 $n\times p$ 矩阵，它的第 i 行，第 k 列元素为

$$\sum_{j=1}^{m}a_{ij}b_{jk}=a_{i1}b_{1k}+a_{i2}b_{2k}+\cdots+a_{im}b_{mk},$$

$1\leqslant i\leqslant n, 1\leqslant k\leqslant p$. 由乘法定义可知它有性质：

(i) 乘法不满足交换律.

(ii) 有乘法结合律，即任给 $n\times m$ 矩阵 A, $m\times p$ 矩阵 B, $p\times q$ 矩阵 C，则有

$$(AB)C=A(BC).$$

(iii) n 阶方阵

$n\times p$ $p\times q$ $n\times m$ $m\times q$
$n\times q$ $n\times q$

$$I^{(n)}=\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix}$$

称为单位方阵，它的对角元素全为 1，其余元素全为零. 它有性质：对任意 $n\times m$ 矩阵 A，则有

$$AI^{(m)}=I^{(n)}A=A.$$

(iv) 乘法不满足消去律，即存在 $n\times m$ 矩阵 $A\neq0$，$m\times p$ 矩阵 $B\neq0$，但是 $AB=0$.

(v) 加、乘分配律，即任取 $n\times m$ 矩阵 A,B 及 $m\times p$ 矩阵 C,D，则有

$$(A+B)C=AC+BC,\quad A(C+D)=AC+AD.$$

(vi) 任取 $n\times m$ 矩阵 A 及 $m\times p$ 矩阵 B，设 λ 为数，则有

$$(\lambda A)B=A(\lambda B)=\lambda(AB)$$

上面一系列性质中，性质(i)及(iv)表明了矩阵乘法和数的相乘，多项式的相乘是有本质差别的.

定义 $n\times m$ 矩阵

$$A=\begin{pmatrix}a_{11} & \cdots & a_{1m}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nm}\end{pmatrix}$$

的共轭矩阵

$$\bar{A}=\begin{pmatrix}\bar{a}_{11} & \cdots & \bar{a}_{1m}\\ \vdots & & \vdots\\ \bar{a}_{n1} & \cdots & \bar{a}_{nm}\end{pmatrix}$$

其中 $\bar{a}_{jk}$ 为 a_{jk} 之共轭复数.

(4) 转置运算和取共轭的运算有性质：

(i) 设 A 和 B 为 $n\times m$ 矩阵，则有

$$(A+B)'=A'+B',\quad (\overline{A+B})=\bar{A}+\bar{B}.$$

(ii) 设 λ 为数, A 为 $n\times m$ 矩阵, 则有

$$(\lambda A)'=\lambda A',\quad \overline{(\lambda A)}=\bar{\lambda}\bar{A}.$$

(iii) 设 A 为 $n\times m$ 矩阵, B 为 $m\times p$ 矩阵, 则有

$$(AB)'=B'A',\quad (\overline{AB})=\bar{A}\bar{B}.$$

(iv) 设 A 为 $n\times m$ 矩阵, 则有

$$(\bar{A}')=(\bar{A})'.$$

下面引进**分块矩阵**，并讨论它的代数运算. 用分块矩阵来计算, 是矩阵论中最基本和重要的计算方法.

设 A 为 $n\times m$ 矩阵、取自然数 $n_1, n_2, \cdots, n_p, m_1, m_2, \cdots, m_q$, 使得

$$\sum_{j=1}^{p} n_j=n,\quad \sum_{k=1}^{q} m_k=m.$$

将矩阵 A 的 n 行, 按照 n_1 行, n_2 行, $\cdots$, n_p 行分开, 将 A 的 m 列, 按照 m_1 列, m_2 列, $\cdots$, m_q 列分开. 便有

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{pmatrix}\begin{matrix} \text{共 } n_1 \text{ 行} \\ \text{共 } n_2 \text{ 行} \\ \vdots \\ \text{共 } n_p \text{ 行} \end{matrix}$$

$$\begin{matrix} \text{共} & \text{共} & & \text{共} \\ m_1 & m_2 & \cdots & m_q \\ \text{列} & \text{列} & & \text{列} \end{matrix}$$

于是 A_{ij} 为 $n_i\times m_j$ 矩阵, 它是 A 的子矩阵.

显然, 同一矩阵可以有各种不同的分块方法. 例如:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}=\begin{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} & \begin{pmatrix} 3 \\ 0 \end{pmatrix} \\ (0 \quad 0) & (1) \end{pmatrix}=\begin{pmatrix} 1 & (2 \quad 3) \\ \begin{pmatrix} 2 \\ 0 \end{pmatrix} & \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix}.$$

下面给出矩阵分块和运算间的关系

(1) 加法

设 A 和 B 都是 $n\times m$ 矩阵，用完全相同的方式分块为

$$A=\begin{pmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pq} \end{pmatrix},\quad B=\begin{pmatrix} B_{11} & \cdots & B_{1q} \\ \vdots & & \vdots \\ B_{p1} & \cdots & B_{pq} \end{pmatrix},$$

则

$$A+B=\begin{pmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pq} \end{pmatrix}+\begin{pmatrix} B_{11} & \cdots & B_{1q} \\ \vdots & & \vdots \\ B_{p1} & \cdots & B_{pq} \end{pmatrix}$$

$$=\begin{pmatrix} A_{11}+B_{11} & \cdots & A_{1q}+B_{1q} \\ \vdots & & \vdots \\ A_{p1}+B_{p1} & \cdots & A_{pq}+B_{pq} \end{pmatrix}.$$

(2) 纯量乘积

$$\lambda A=\lambda\begin{pmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pq} \end{pmatrix}=\begin{pmatrix} \lambda A_{11} & \cdots & \lambda A_{1q} \\ \vdots & & \vdots \\ \lambda A_{p1} & \cdots & \lambda A_{pq} \end{pmatrix},$$

(3) 乘法

设 A 为 $n\times m$ 矩阵，按照 $n_1, n_2, \cdots, n_p$ 行，$m_1, m_2, \cdots, m_q$ 列分块为

$$A=\begin{pmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pq} \end{pmatrix}\begin{matrix} n_1 \\ \vdots \\ n_p \end{matrix}$$
$$\begin{matrix} m_1 & \cdots & m_q \end{matrix}$$

设 B 为 $m\times f$ 矩阵，对 B 作分块的方式不能任意，要求对 B 的行的分块方式和对 A 的列的分块方式相同，即

$$B=\begin{pmatrix} B_{11} & \cdots & B_{1r} \\ \vdots & & \vdots \\ B_{q1} & \cdots & B_{qr} \end{pmatrix}\begin{matrix} m_1 \\ \vdots \\ m_q \end{matrix}$$
$$\begin{matrix} s_1 & \cdots & s_r \end{matrix}$$

所以 A_{ij} 为 $n_i\times m_j$ 矩阵，B_{jk} 为 $m_j\times s_k$ 矩阵. 因此 A_{ij} 和 B_{jk} 可以相乘. 且

$$AB=\begin{pmatrix}A_{11} & \cdots & A_{1q}\\ \vdots & & \vdots\\ A_{p1} & \cdots & A_{pq}\end{pmatrix}\begin{pmatrix}B_{11} & \cdots & B_{1r}\\ \vdots & & \vdots\\ B_{q1} & \cdots & B_{qr}\end{pmatrix}$$

$$=\begin{pmatrix}\sum_{j=1}^{q}A_{1j}B_{j1} & \cdots & \sum_{j=1}^{q}A_{1j}B_{jr}\\ \vdots & & \vdots\\ \sum_{j=1}^{q}A_{pj}B_{j1} & \cdots & \sum_{j=1}^{q}A_{pj}B_{jr}\end{pmatrix}.$$

事实上

$$AB=\begin{pmatrix}a_{11} & \cdots & a_{1m}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nm}\end{pmatrix}\begin{pmatrix}b_{11} & \cdots & b_{1f}\\ \vdots & & \vdots\\ b_{m1} & \cdots & b_{mf}\end{pmatrix}$$

$$=\begin{pmatrix}\sum_{j=1}^{m}a_{1j}b_{j1} & \cdots & \sum_{j=1}^{m}a_{1j}b_{jf}\\ \vdots & & \vdots\\ \sum_{j=1}^{m}a_{nj}b_{j1} & \cdots & \sum_{j=1}^{m}a_{nj}b_{jf}\end{pmatrix},$$

它可改写为

$$AB=\begin{pmatrix}(a_{11} & \cdots & a_{1m}) & B\\ & \vdots & & \\ (a_{n1} & \cdots & a_{nm}) & B\end{pmatrix}=\begin{pmatrix}\begin{pmatrix}a_{11} & \cdots & a_{1m}\\ \vdots & & \vdots\\ a_{n_11} & \cdots & a_{n_1m}\end{pmatrix}B\\ \cdots\\ \begin{pmatrix}a_{n'1} & \cdots & a_{n'm}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nm}\end{pmatrix}B\end{pmatrix}$$

$$=\begin{pmatrix}(A_{11} & \cdots & A_{1q})\ B\\ & \vdots & \\ (A_{p1} & \cdots & A_{pq})\ B\end{pmatrix},$$

其中 $n'=n_1+\cdots+n_{p-1}+1$. 同理

$$AB=\begin{pmatrix}(A_{11}\cdots A_{1q})\begin{pmatrix}B_{11}\\ \vdots\\ B_{q1}\end{pmatrix} & \cdots & (A_{11}\cdots A_{1q})\begin{pmatrix}B_{1r}\\ \vdots\\ B_{qr}\end{pmatrix}\\ \vdots & & \vdots\\ (A_{p1}\cdots A_{pq})\begin{pmatrix}B_{11}\\ \vdots\\ B_{q_1}\end{pmatrix} & \cdots & (A_{p1}\cdots A_{pq})\begin{pmatrix}B_{1r}\\ \vdots\\ B_{qr}\end{pmatrix}\end{pmatrix}$$

因此问题化为证明

$$(A_{i1}\ \cdots\ A_{iq})\begin{pmatrix}B_{1l}\\ \vdots\\ B_{ql}\end{pmatrix}=\sum_{j=1}^{q}A_{ij}B_{jl}.$$

事实上，若能证

$$(C^{(n,m_1)}\ D^{(n,m_2)})\begin{pmatrix}E^{(m_1,f)}\\ F^{(m_2,f)}\end{pmatrix}=CE+DF,$$

则

$$(A_{i1}\cdots A_{iq})\begin{pmatrix}B_{1l}\\ \vdots\\ B_{ql}\end{pmatrix}=(A_{i1}(A_{i2}\cdots A_{iq}))\begin{pmatrix}B_{1l}\\ \begin{pmatrix}B_{2l}\\ \vdots\\ B_{ql}\end{pmatrix}\end{pmatrix}$$

$$=A_{i1}B_{1l}+(A_{i2}\cdots A_{iq})\begin{pmatrix}B_{2l}\\ \vdots\\ B_{ql}\end{pmatrix}$$

$$=\cdots=\sum_{j=1}^{q}A_{ij}B_{jl}.$$

下面来证 $(C \quad D)\begin{pmatrix}E\\F\end{pmatrix}=CE+DF$. 取出 C 的第 i 行 $(c_{i1}\cdots c_{im_1})$，D 的第 i 行 $(d_{i1},\cdots,d_{im_2})$；取出 E 的第 j 列，排成横行为 $(e_{1j},\cdots,e_{m_1j})$，F 的第 j 列，排成横行为 $(f_{1j},\cdots,f_{m_2j})$. 于是 $CE+DF$ 的第 i 行，第 j 列元素为

$$\sum_{k=1}^{m_1} c_{ik}e_{kj}+\sum_{k=1}^{m_2} d_{ik}f_{kj}.$$

而 $(C\ D)$ 的第 i 行为 $(c_{i1},\cdots,c_{im_1},d_{i1},\cdots,d_{im_2})$，$\begin{pmatrix}E\\F\end{pmatrix}$ 的第 j 列排成横行为 $(e_{1j},\cdots,e_{m_1j},f_{1j},\cdots,f_{m_2j})$，于是 $(C \quad D)\begin{pmatrix}E\\F\end{pmatrix}$ 的第 i 行，第 j 列元素为

$$c_{i1}e_{1j}+\cdots+c_{im_1}e_{m_1j}+d_{i1}f_{1j}+\cdots+d_{im_2}f_{m_2j}.$$

这证明了 $CE+DF$ 和 $(C\ D)\begin{pmatrix}E\\F\end{pmatrix}$ 的任意相同位置元素相等，即 $(C \quad D)\begin{pmatrix}E\\F\end{pmatrix}=CE+DF$. 所以完全证明了分块矩阵的乘法公式.

由乘法公式可知，乘积仍为分块矩阵，其行的分块方式和 A 的行的分块方式相同，列的分块方式和 B 的列的分块方式相同. 而且乘法公式也告诉我们，对分块矩阵作乘法，相当于把每一块当作一个元素，按通常矩阵乘法的定义作运算. 这一特点是使用分块矩阵的主要原因.

(4) 转置运算

$$A'=\begin{pmatrix}A_{11} & \cdots & A_{1q}\\ \vdots & & \vdots\\ A_{p1} & \cdots & A_{pq}\end{pmatrix}'=\begin{pmatrix}A'_{11} & \cdots & A'_{p1}\\ \vdots & & \vdots\\ A'_{1q} & \cdots & A'_{pq}\end{pmatrix},$$

(5) 共轭运算

$$\bar{A}=\overline{\begin{pmatrix}A_{11} & \cdots & A_{1q}\\ \vdots & & \vdots\\ A_{p1} & \cdots & A_{pq}\end{pmatrix}}=\begin{pmatrix}\bar{A}_{11} & \cdots & \bar{A}_{1q}\\ \vdots & & \vdots\\ \bar{A}_{p1} & \cdots & \bar{A}_{pq}\end{pmatrix}.$$

定义　n 阶方阵 $A=(a_{ij})$ 称为**对角方阵**，如果 $a_{ij}=0, i\neq j$. 这时可记作

$$A=\text{diag}(a_{11},\cdots,a_{nn}).$$

n 阶方阵 $A=(a_{ij})$ 称为**上三角方阵**，如果 $a_{ij}=0, i>j$；称为**下三角方阵**．如果 $a_{ij}=0, i<j$. n 阶方阵 A 称为**对称方阵**，如果 $a_{ij}=a_{ji}, 1\leqslant i, j\leqslant n$，即有 $A'=A$；称为**斜对称方阵**，如果 $a_{ji}=-a_{ij}$, $1\leqslant i, j\leqslant n$，即有 $A'=-A$；称为**Hermite 方阵**，如果 $a_{ij}=\bar{a}_{ji}, 1\leqslant i, j\leqslant n$，即有 $\bar{A}'=A$；称为**斜 Hermite 方阵**，如果 $a_{ij}=-\bar{a}_{ji}$, $1\leqslant i, j\leqslant n$，即有 $\bar{A}'=-A$.

对角方阵，对称方阵，斜对称方阵，Hermite 方阵及斜 Hermite 方阵是以后常要见到的方阵.

定义　$n\times m$ 矩阵 A 分块为

$$A=\begin{pmatrix}A_{11} & \cdots & A_{1p}\\ \vdots & & \vdots\\ A_{p1} & \cdots & A_{pp}\end{pmatrix}$$

则当 $A_{ij}=0$，$i\neq j$ 时，A 称为**准对角方阵**，它可改记为:

$$A=\text{diag}(A_{11},\cdots,A_{pp}).$$

(这里要注意尽管行和列的块数相同，但是 A_{ii} 不一定是方阵)，又当 $A_{ij}=0, i>j$ 时，则 A 称为**准上三角方阵**；当 $A_{ij}=0, i<j$ 时，则 A 称为**准下三角方阵**.

注意"准"的概念和分块方式有关．例如下面一个例子，用三种不同方式分块，则它可以分别为准对角，准上三角和准下三角．记

$$A=\begin{pmatrix}1&2&0&0\\3&8&0&0\\0&0&1&3\\0&0&2&1\\0&0&0&7\end{pmatrix}.$$

三种分块方式分别为

$$\begin{pmatrix}\begin{pmatrix}1&2\\3&8\end{pmatrix}&\begin{pmatrix}0&0\\0&0\end{pmatrix}\\\begin{pmatrix}0&0\\0&0\\0&0\end{pmatrix}&\begin{pmatrix}1&3\\2&1\\0&7\end{pmatrix}\end{pmatrix},\quad\begin{pmatrix}\begin{pmatrix}1&2\\3&8\\0&0\end{pmatrix}&\begin{pmatrix}0&0\\0&0\\1&3\end{pmatrix}\\\begin{pmatrix}0&0\\0&0\end{pmatrix}&\begin{pmatrix}2&1\\0&7\end{pmatrix}\end{pmatrix},\quad\begin{pmatrix}(1\ \ 2)&(0\ \ 0)\\\begin{pmatrix}3&8\\0&0\\0&0\\0&0\end{pmatrix}&\begin{pmatrix}0&0\\1&3\\2&1\\0&7\end{pmatrix}\end{pmatrix}.$$

习题 4.1

1. 试证：对任一数 λ 及任一 n 阶方阵 A，

$$\det(\lambda A)=\lambda^n\det A.$$

2. 试各举一例说明存在 n 阶方阵 A,B，使得

(i) $AB\neq BA$，

(ii) $A\neq 0, B\neq 0, AB=0$，

(iii) $\det(A+B)\neq\det A+\det B$.

3. 试求下列各矩阵的乘积：

(i) $\mathrm{diag}(\lambda_1,\cdots,\lambda_n)^p$，

(ii) $\begin{pmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{pmatrix}^n$，

(iii) $\begin{pmatrix}1&\dfrac{a}{n}\\-\dfrac{a}{n}&1\end{pmatrix}^n$，

(iv) $\begin{pmatrix}\lambda&1&&&\\&\lambda&1&&\\&&\ddots&\ddots&\\&&&\lambda&1\\&&&&\lambda\end{pmatrix}^p$，其中方阵为 n 阶的.

4. 取定 n 阶方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

记 a_{ij} 的代数余子式为 A_{ij}, $1\leqslant i,j\leqslant n$, 作 n 阶方阵

$$A^*=\begin{pmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{nn} \end{pmatrix}$$

则 A^* 称为 A 的**伴随方阵**. 试证: 对 n 阶方阵 A, 有

$$AA^*=A^*A=(\det A)I^{(n)}.$$

5. 设 n 阶方阵 A 和 B 可交换, 即有 $AB=BA$. 试证:

(i) $(A+B)^m=\sum_{k=0}^{m}\binom{m}{k}A^kB^{m-k}$,

(ii) $A^3-B^3=(A-B)(A^2+AB+B^2)$.

6. 设 A 为 n 阶上三角方阵, 且对角元素都等于零, 试证: A 为**幂零方阵**, 即 $A^n=0$.

7. 设 $A_1, A_2, \cdots, A_n$ 都是上三角方阵, 且对角元素都等于零. 试证: $A_1A_2\cdots A_n=0$.

8. 试证: 与任一 n 阶方阵可交换的方阵都是**纯量方阵**, 即为 $\lambda I^{(n)}$, 其中 λ 为数, I 为 n 阶单位方阵.

9. 给定 s 个不同的数 $\lambda_1, \lambda_2, \cdots, \lambda_s$, 记 $I_1, I_2, \cdots, I_s$ 分别为 $n_1, n_2, \cdots, n_s$ 阶单位方阵, 记 $\sum_{j=1}^{s}n_j=n$. 试证: 和 n 阶对角方阵 $\mathrm{diag}(\lambda_1I_1, \cdots, \lambda_sI_s)$ 可交换的方阵必为

$$\mathrm{diag}(B_1, \cdots, B_s)$$

其中 $B_1, \cdots, B_s$ 分别为 $n_1, \cdots, n_s$ 阶方阵.

§4.2 Binet-Cauchy 公式

定理 4.2.1 (Binet-Cauchy 公式) 设 A 和 B 分别为 $p\times q$ 和 $q\times p$ 矩阵, 则有

$$\det AB=\begin{cases}0, & \text{当 } p>q \text{ 时},\\ \det A\cdot\det B, & \text{当 } p=q \text{ 时},\\ \sum\limits_{1\leqslant i_1<\cdots<i_p\leqslant q} A\begin{pmatrix}1 & 2 & \cdots & p\\ i_1 & i_2 & \cdots & i_p\end{pmatrix} B\begin{pmatrix}i_1 & i_2 & \cdots & i_p\\ 1 & 2 & \cdots & p\end{pmatrix}, & \text{当 } p<q \text{ 时}.\end{cases}$$

证 记

$$A=\begin{pmatrix}a_{11} & \cdots & a_{1q}\\ \vdots & & \vdots\\ a_{p1} & \cdots & a_{pq}\end{pmatrix},\quad B=\begin{pmatrix}b_{11} & \cdots & b_{1p}\\ \vdots & & \vdots\\ b_{q1} & \cdots & b_{qp}\end{pmatrix}.$$

则

$$\det AB=\det\begin{pmatrix}\sum\limits_{j_1=1}^{q}a_{1j_1}b_{j_11} & \cdots & \sum\limits_{j_p=1}^{q}a_{1j_p}b_{j_pp}\\ \vdots & & \vdots\\ \sum\limits_{j_1=1}^{q}a_{pj_1}b_{j_11} & \cdots & \sum\limits_{j_p=1}^{q}a_{pj_p}b_{j_pp}\end{pmatrix}$$

$$=\sum_{j_1,\cdots,j_p=1}^{q} b_{j_11}b_{j_22}\cdots b_{j_pp}\det\begin{pmatrix}a_{1j_1} & \cdots & a_{1j_p}\\ \vdots & & \vdots\\ a_{pj_1} & \cdots & a_{pj_p}\end{pmatrix}$$

当 $p>q$ 时，$j_1,j_2,\cdots,j_p\in\{1,2,\cdots,q\}$，所以行列式中总有两列相同，因此 $\det AB=0$. 下面设 $p\leqslant q$，这时

$$\det AB=\sum_{\substack{j_1,\cdots,j_p=1\\ j_1,\cdots,j_p\text{两两不等}}}^{q} b_{j_11}b_{j_22}\cdots b_{j_pp}\det\begin{pmatrix}a_{1j_1} & \cdots & a_{1j_p}\\ \vdots & & \vdots\\ a_{pj_1} & \cdots & a_{pj_p}\end{pmatrix}$$

由引理 2.1.9，有

$$\det AB=\sum_{1\leqslant i_1<\cdots<i_p\leqslant q}\ \sum_{\binom{i_1 i_2\cdots i_p}{j_1 j_2\cdots j_p}} b_{j_11}b_{j_22}\cdots b_{j_pp}\det\begin{pmatrix}a_{1j_1} & \cdots & a_{1j_p}\\ \vdots & & \vdots\\ a_{pj_1} & \cdots & a_{pj_p}\end{pmatrix}$$

$$=\sum_{1\leqslant i_1<\cdots<i_p\leqslant q}\ \sum_{\binom{i_1 i_2\cdots i_p}{j_1 j_2\cdots j_p}} \delta^{i_1 i_2\cdots i_p}_{j_1 j_2\cdots j_p}$$

$$\cdot b_{j_1 1} b_{j_2 2} \cdots b_{j_p p} \det\begin{pmatrix} a_{1i_1} & \cdots & a_{1i_p} \\ \vdots & & \vdots \\ a_{pi_1} & \cdots & a_{pi_p} \end{pmatrix}$$

由行列式之定义可知

$$\det AB = \sum_{1 \leqslant i_1 < \cdots < i_p \leqslant q} A\binom{1\,2\,\cdots\,p}{i_1\,i_2\cdots i_p} B\binom{i_1\,i_2\cdots i_p}{1\,2\,\cdots\,p}.$$

所以当 $p<q$ 时定理成立. 当 $q=p$ 时, $i_1=1$, $\cdots$, $i_p=p$, 于是 $\det AB=\det A\ \det B$. 证完.

定理4.2.2 设 A 和 B 分别为 $p\times q$ 和 $q\times s$ 矩阵, 则 $p\times s$ 矩阵 $C=AB$ 的 r 阶子式

$$C\binom{i_1\,i_2\cdots i_r}{j_1\,j_2\cdots j_r} = \begin{cases} 0, & \text{当 } r>q \text{ 时;} \\ \displaystyle\sum_{1\leqslant k_1<\cdots\cdots<k_r\leqslant q} A\binom{i_1\,i_2\cdots i_r}{k_1\,k_2\cdots k_r} B\binom{k_1\,k_2\cdots k_r}{j_1\,j_2\cdots j_r}, & \text{当 } r\leqslant q \text{ 时,} \end{cases}$$

其中 $1\leqslant i_1<\cdots<i_r\leqslant p$, $1\leqslant j_1<\cdots<j_r\leqslant s$.

证 记

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1q} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pq} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & \cdots & b_{1s} \\ \vdots & & \vdots \\ b_{q1} & \cdots & b_{qs} \end{pmatrix},$$

则

$$C=AB=\begin{pmatrix} \sum_{k=1}^{q} a_{1k}b_{k1} & \cdots & \sum_{k=1}^{q} a_{1k}b_{ks} \\ \vdots & & \vdots \\ \sum_{k=1}^{q} a_{pk}b_{k1} & \cdots & \sum_{k=1}^{q} a_{pk}b_{ks} \end{pmatrix}.$$

于是

$$C\binom{i_1\cdots i_r}{j_1\cdots j_r} = \det\begin{pmatrix} \sum_{k=1}^{q} a_{i_1k}b_{kj_1} & \cdots & \sum_{k=1}^{q} a_{i_1k}b_{kj_r} \\ \vdots & & \vdots \\ \sum_{k=1}^{q} a_{i_rk}b_{kj_1} & \cdots & \sum_{k=1}^{q} a_{i_rk}b_{kj_r} \end{pmatrix}.$$

$$=\det\begin{pmatrix} a_{i_1 1} & \cdots & a_{i_1 q} \\ \vdots & & \vdots \\ a_{i_r 1} & \cdots & a_{i_r q}\end{pmatrix}\begin{pmatrix} b_{1j_1} & \cdots & b_{1j_r} \\ \vdots & & \vdots \\ b_{qj_1} & \cdots & b_{qj_r}\end{pmatrix}.$$

由 Binet-Cauchy公式，便证明了定理．证完．

Binet-Cauchy 公式中最有用的情形为：当 A 和 B 都是 n 阶方阵的情形，这时有

$$\det AB=\det A\cdot\det B.$$

用这个公式可以计算行列式．例如

例 试求 n 阶轮回方阵

$$A=\begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \cdots & a_0\end{pmatrix}$$

的行列式．

解 记

$$\omega=\cos\frac{2\pi}{n}+\sqrt{-1}\sin\frac{2\pi}{n}=e^{\frac{2\pi\sqrt{-1}}{n}}.$$

作 n 阶 Vandermonde 方阵

$$V=V(1,\omega,\omega^2,\cdots,\omega^{n-1})=\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)(n-1)}\end{pmatrix}.$$

于是

$$AV=\begin{pmatrix}\sum_{k=0}^{n-1}a_k & \sum_{k=0}^{n-1}a_k\omega^k & \cdots & \sum_{k=0}^{n-1}a_k\omega^{k(n-1)}\\ \sum_{k=0}^{n-1}a_k & \omega\sum_{k=0}^{n-1}a_k\omega^k & \cdots & \omega^{n-1}\sum_{k=0}^{n-1}a_k\omega^{k(n-1)}\\ \vdots & \vdots & & \vdots\\ \sum_{k=0}^{n-1}a_k & \omega^{n-1}\sum_{k=0}^{n-1}a_k\omega^k & \cdots & \omega^{(n-1)^2}\sum_{k=0}^{n-1}a_k\omega^{k(n-1)}\end{pmatrix}.$$

记

$$f(x)=a_0+a_1x+\cdots+a_{n-}x^{n-1},$$

于是

$$f(\omega^j)=a_0\omega^{0j}+a_1\omega^{1j}+\cdots+a_{n-1}\omega^{j(n-1)}=\sum_{k=0}^{n-1}a_k\omega^{kj}.$$

因此

$$AV=\begin{pmatrix}f_1(\omega^0) & f(\omega) & \cdots & f(\omega^{n-1})\\ f(\omega^0) & \omega f(\omega) & \cdots & \omega^{n-1}f(\omega^{n-1})\\ \vdots & \vdots & & \vdots\\ f(\omega^0) & \omega^{n-1}f(\omega) & \cdots & \omega^{(n-)^2}f(\omega^{n-1})\end{pmatrix}$$

$$=V\mathrm{diag}(f(\omega^0),f(\omega),\cdots,f(\omega^{n-1})).$$

双方取行列式，有

$$\det A\ \det V=\det V\cdot\prod_{j=0}^{n-1}f(\omega^j),$$

由于 $\det V\neq 0$，所以证明了

$$\det A=\prod_{j=0}^{n-1}f(\omega^j).$$

习题 4.2

1. 试利用 Laplace 展开式来证明：设 A 和 B 为 n 阶方阵，I 为 n 阶单位方阵，则有

$$\det\begin{pmatrix} A & 0 \\ -I & B \end{pmatrix}=\det A\det B.$$

再用行列式的性质来证明：

$$\det\begin{pmatrix} A & 0 \\ -I & B \end{pmatrix}=\det AB \quad .$$

从而导出公式 $\det AB=\det A\ \det B$.

2. 设 $i<j, k<l$. 记 n 阶方阵 A 的元素 $a_{ik}, a_{il}, a_{jk}, a_{jl}$ 的代数余子式分别为 $A_{ik}, A_{il}, A_{jk}, A_{jl}$. 试证：

$$\det\begin{pmatrix} A_{ik} & A_{il} \\ A_{jk} & A_{jl} \end{pmatrix}=(-1)^{i+j+k+l} \quad A\begin{pmatrix} 1\cdots(i-1)(i+1)\cdots(j-1)(j+1)\cdots n \\ 1\cdots(k-1)(k+1)\cdots(l-1)(l+1)\cdots n \end{pmatrix}\det A.$$

3. 计算下面 n 阶行列式

$$\text{(i)}\ \det\begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ \mu a_{n-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & & \vdots \\ \mu a_1 & \mu a_2 & \cdots & a_0 \end{pmatrix} \qquad \text{(ii)}\ \det\begin{pmatrix} s_0 & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ \vdots & \vdots & & \vdots \\ s_{n-1} & s_n & \cdots & s_{2n-2} \end{pmatrix},$$

其中 $s_j=\sum_{l=1}^{n} x_l^j$, $x_1, \cdots, x_n$ 为 n 个数.

4. 记 A^* 为 n 阶方阵 A 的伴随方阵(见习题 4.1 第 4 题)，试证：

(i) $\det A^*=(\det A)^{n-1}$,

(ii) 进一步有，任取 $1\leqslant i_1<\cdots<i_p\leqslant n$, $1\leqslant i_{p+1}<\cdots<i_n\leqslant n$, $1\leqslant j_1<\cdots<j_p\leqslant m$, $1\leqslant j_{p+1}<\cdots<j_n\leqslant n$, 其中 $i_1 i_2\cdots i_n$ 及 $j_1 j_2\cdots j_n$ 都是 $1, 2, \cdots, n$ 的排列，试证：

$$A^*\begin{pmatrix} i_1 i_2\cdots i_p \\ j_1 j_2\cdots j_p \end{pmatrix}=(\det A)^{p-1}\delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n} A\begin{pmatrix} j_{p+1} j_{p+2}\cdots j_n \\ i_{p+1} i_{p+2}\cdots i_n \end{pmatrix}.$$

5. 设 A 为 $2\times n$ 实矩阵. 试分别由定义及 Binet-Cauchy 公式来计算 $\det AA'$，从而推出 **Cauchy不等式**. 设 A 为 $2\times n$ 复矩阵，试分别由定义及 Binet-Cauchy 公式来计算 $\det A\bar{A}'$，从而推出 Cauchy 不等式.

6. 利用 Binet-Cauchy 公式及 Cauchy 不等式，试证：对任两 $p\times q$ 实矩阵 A 和 B，有

$$\det AA'\ \det BB'\geqslant(\det AB')^2.$$

7. 设 A 为 $p\times q$ 复矩阵，试证：p 阶方阵 $A\bar{A}'$ 的任一主子式 $(A\bar{A}')\begin{pmatrix} i_1 i_2\cdots i_r \\ i_1 i_2\cdots i_r \end{pmatrix}\geqslant 0$, $1\leqslant i_1<\cdots<i_r\leqslant p$.

8. 试证：

$$\det\begin{pmatrix} x_1 & -x_2 & -x_3 & -x_4 \\ x_2 & x_1 & -x_4 & x_3 \\ x_3 & x_4 & x_1 & -x_2 \\ x_4 & -x_3 & x_2 & x_1 \end{pmatrix}=(x_1^2+x_2^2+x_3^2+x_4^2)^2.$$

9. 设A和B为n阶方阵，试证：AB和BA的p阶主子式之和必相等，$1\leqslant p\leqslant n$. 从而证明：

$$\det(\lambda I^{(n)}-AB)=\det(\lambda I^{(n)}-BA)$$

对任意数λ成立. 特别$\det(I-AB)=\det(I-BA)$.

10. 设A和B分别为$n\times m$及$m\times n$矩阵. 记

$$C=\begin{pmatrix} \lambda I^{(n)} & 0 \\ 0 & I^{(m)} \end{pmatrix}-\begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}$$

试证：存在$n+m$阶方阵P适合

(1) $\det PC=\lambda^m\det(\lambda I^{(n)}-AB)$,

(2) $\det CP=\lambda^n\det(\lambda I^{(m)}-BA)$.

由此可推出

$$\lambda^m\det(\lambda I^{(n)}-AB)=\lambda^n\det(\lambda I^{(m)}-BA).$$

§4.3 逆 方 阵

定义 n阶方阵的迹为对角元素之和，记作tr. 即设

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

则

$$\mathrm{tr}A=a_{11}+a_{22}+\cdots+a_{nn}.$$

由定义可知，对任意的n阶方阵A,B及数λ，则有

(i) $\mathrm{tr}(A+B)=\mathrm{tr}A+\mathrm{tr}B$,

(ii) $\mathrm{tr}(\lambda A)=\lambda\mathrm{tr}A$,

(iii) $\mathrm{tr}(AB)=\mathrm{tr}BA$,

(iv) $\mathrm{tr}A'=\mathrm{tr}A$,

(v) $\mathrm{tr}\,\overline{A}=\overline{(\mathrm{tr}A)}$,

(vi) $\operatorname{tr}A\bar{A}'=0$ 当且仅当 $A=0$.

定义 设 A 为 n 阶方阵，$I^{(n)}$ 为 n 阶单位方阵，如果存在 n 阶方阵 B，使得

$$AB=BA=I,$$

则 A 称为**可逆方阵**，否则称为**不可逆方阵**. 当 A 为可逆方阵时，B 可改记为 A^{-1}，称为 A 的**逆方阵**.

定义 设 A 为 n 阶方阵，则当 $\det A=0$ 时 A 称为**奇异方阵**；当 $\det A\neq 0$ 时 A 称为**非异方阵**.

定理 4.3.1 n 阶方阵 A 可逆当且仅当非异. 且可逆方阵的逆方阵唯一存在，它是

$$A^{-1}=(\det A)^{-1}A^*,$$

其中 A^* 为 A 的伴随方阵.

证 由 Laplace 展开定理可知

$$AA^*=A^*A=(\det A)I.$$

所以若 $\det A\neq 0$，令 $B=(\det A)^{-1}A^*$ 便有 $AB=BA=I$. 反之，若存在 B 使得 $AB=BA=I$，则 $\det A\det B=1$，这证明了 $\det A\neq 0$. 最后证唯一性. 今若有 $AC=CA=I$，其中 C 为 n 阶方阵，取 $B=(\det A)^{-1}A^*$，则

$$B=B(AC)=(BA)C=I\cdot C=C.$$

这证明了 $C=B$，即 A 的逆方阵唯一. 证完.

由定义可知，逆方阵有如下性质：设 A 和 B 都是 n 阶可逆方阵，λ 为非零常数，则有

(1) $(A^{-1})^{-1}=A$,

(2) $(AB)^{-1}=B^{-1}A^{-1}$，特别 $(\lambda A)^{-1}=\lambda^{-1}A^{-1}$,

(3) $(A')^{-1}=(A^{-1})'$,

(4) $(\bar{A})^{-1}=\overline{(A^{-1})}$.

利用逆方阵还可以给出 Cramer 法则一个简单证明. 首先,

利用矩阵乘法可将线性方程组写成简单的形式，记

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \beta=\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

于是线性方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+\cdots+a_{nn}x_n=b_n. \end{cases}$$

可改写为

$$Ax=\beta.$$

设 $\det A\neq 0$，则有解

$$x=(A^{-1}A)x=A^{-1}(Ax)=A^{-1}\beta.$$

由于 $A^{-1}=(\det A)^{-1}A^*$，于是

$$x_i=\left(\sum_{j=1}^{n}A_{ji}b_j\right)\Big/\det A, \quad i=1,2,\cdots,n.$$

这就是定理 2.5.1 的结论.

最后，我们介绍矩阵计算中的最基本技巧——打洞.

设 A 为 $n\times m$ 矩阵，将它分成四块为

$$A=\begin{pmatrix} B^{(r)} & C^{(r,m-r)} \\ D^{(n-r,r)} & E^{(n-r,m-r)} \end{pmatrix}.$$

于是当 $\det B\neq 0$ 时有 Schur 公式

$$\begin{pmatrix} I^{(r)} & 0 \\ -DB^{-1} & I \end{pmatrix}\begin{pmatrix} B & C \\ D & E \end{pmatrix}=\begin{pmatrix} B & C \\ 0 & E-DB^{-1}C \end{pmatrix},$$

$$\begin{pmatrix} B & C \\ D & E \end{pmatrix}\begin{pmatrix} I^{(r)} & -B^{-1}C \\ 0 & I \end{pmatrix}=\begin{pmatrix} B & 0 \\ D & E-DB^{-1}C \end{pmatrix}.$$

用**打洞技巧**，可以将高阶行列式的计算化为低阶行列式的计算.

定理 4.3.2 设 A 为 n 阶方阵，按照 $r, n-r$ 行、列作分块，记作

$$A=\begin{pmatrix} B & C \\ D & E \end{pmatrix}.$$

设 $\det B \neq 0$，则有

$$\begin{pmatrix} B & C \\ D & E \end{pmatrix}=\begin{pmatrix} I & 0 \\ DB^{-1} & I \end{pmatrix}\begin{pmatrix} B & 0 \\ 0 & E-DB^{-1}C \end{pmatrix}\begin{pmatrix} I & B^{-1}C \\ 0 & I \end{pmatrix}$$

所以有

$$\det A=\det B \det(E-DB^{-1}C).$$

因此当 A 及 B 都可逆时，$E-DB^{-1}C$ 也可逆. 且

$$\begin{aligned} A^{-1}&=\begin{pmatrix} B & C \\ D & E \end{pmatrix}^{-1} \\ &=\begin{pmatrix} I & -B^{-1}C \\ 0 & I \end{pmatrix}\begin{pmatrix} B^{-1} & 0 \\ 0 & (E-DB^{-1}C)^{-1} \end{pmatrix}\begin{pmatrix} I & 0 \\ -DB^{-1} & I \end{pmatrix}. \end{aligned}$$

下面给出另一种矩阵技巧，即所谓**摄动法**.

例 将 $2n$ 阶方阵 A 按 n，n 行列分块为

$$A=\begin{pmatrix} B & C \\ D & E \end{pmatrix}.$$

则当 $BC=CB$ 时有

$$\det A=\det(EB-DC).$$

证 引进实参数 t，记 $B_t=tI+B$. 于是$\det B_t=\det(tI+B)$为 t 的 n 次多项式，首项系数为 1，因此它至多有 n 个不同根. 于是存在序列 $t_1, t_2, \cdots$ 使得它们都不是 $\det B_t$ 之根，且有 $t_j \to 0$. 考虑 $2n$ 阶方阵

$$A_t=\begin{pmatrix} B_t & C \\ D & E \end{pmatrix},$$

由 $B_t=tI+B$ 及$BC=CB$ 可知 $B_tC=CB_t$. 取 $t=t_j$，则有$\det B_{t_j}\neq$

0，于是有 $CB_{t_j}^{-1}=B_{t_j}^{-1}C$. 而由定理 4.3.2,

$$\begin{aligned}\det A_{t_j}&=\det B_{t_j}\det(E-DB_{t_j}^{-1}C)\\&=\det B_{t_j}\det(E-DCB_{t_j}^{-1})=\det(E-DCB_{t_j}^{-1})B_{t_j}\\&=\det(EB_{t_j}-DC)=\det(EB-DC+t_jE).\end{aligned}$$

两边关于 t_j 都是多项式，取 $t_j\to 0$，便证明了

$$\det A=\det(EB-DC).$$

证完.

习题 4.3

1. 试举一例说明：存在 n 阶可逆方阵 A,B，使得 $A+B$ 也可逆，但是 $(A+B)^{-1}\neq A^{-1}+B^{-1}$.

2. 试证：不存在 n 阶方阵 A,B，使得 $AB=I+BA$，其中 I 为单位方阵.

3. 设 λ 为数，A,B 为 $n\times 1$ 矩阵，试证：

$$\det(I^{(n)}-\lambda AB')=1-\lambda B'A.$$

什么时候 $I^{(n)}-\lambda AB'$ 有逆，逆是什么？

4. 设 n 阶方阵

$$A=\begin{pmatrix}P^{(r)} & Q\\ 0 & I\end{pmatrix}$$

有 $\det(I-P)\neq 0$，试证：

$$A^m=\begin{pmatrix}P^m & (P^m-I)(P-I)^{-1}Q\\ 0 & I\end{pmatrix},\quad m=0,1,2,\cdots$$

当 $\det P\neq 0$，试证：上式对 $m=-1,-2,\cdots$ 也对.

5. 设 A 为 n 阶幂零方阵，即存在自然数 N，使得 $A^N=0$. 试证：$I-A$ 非异，且

$$(I-A)^{-1}=I+A+A^2+\cdots+A^{N-1}.$$

6. 试求 n 阶 Vandermonde 方阵 $\widetilde{V}(1,\omega,\omega^2,\cdots,\omega^{n-1})$ 的逆方阵，其中 $\omega=\cos\frac{2\pi}{n}+\sqrt{-1}\sin\frac{2\pi}{n}=\exp\frac{2\pi\sqrt{-1}}{n}$，再求 n 阶方阵 A 的逆方阵，这里

$$A=\begin{pmatrix}I^{(r)} & P^{(r,n-r)}\\ Q^{(n-r,r)} & I^{(n-r)}\end{pmatrix},$$

$$\widetilde{V}=\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)^2} \end{pmatrix},$$

又 $I-QP$ 可逆.

7. 设 A 和 B 分别为 $n\times m$ 及 $m\times n$ 矩阵,试证: $I-AB$ 非异当且仅当 $I-BA$ 非异. 试用 A,B 及 $I-AB$ 的逆方阵来表达 $I-BA$ 的逆方阵.

8. 设 A 和 B 为 n 阶方阵,$A^*,B^*,(AB)^*$ 分别为 A,B, AB 的伴随方阵,试证:

$$(AB)^*=B^*A^*.$$

9. 试证:$n+m$ 阶方阵

$$\begin{pmatrix} B^{(n,m)} & I^{(n)} \\ 0 & A^{(m,n)} \end{pmatrix}$$

非异当且仅当 m 阶方阵 AB 可逆.

§4.4 初等变换和矩阵的相抵

下面给出三类 n 阶非异方阵,统称为**初等方阵**. 它们是

$$P_{jk}=\begin{pmatrix} I^{(j-1)} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & I^{(k-j-1)} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I^{(n-k)} \end{pmatrix}$$

$$=\sum_{l\neq j,k} E_{ll}+E_{jk}+E_{kj},\ 1\leqslant j<k\leqslant n$$

称为置换方阵,或称为第一类初等方阵.

一般,一批置换方阵的乘积也称为置换方阵.

$$P_j(a)=\begin{pmatrix} I^{(j-1)} & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & I^{(n-j)} \end{pmatrix}$$

$$=\sum_{l\neq j} E_{ll}+aE_{jj},\quad 1\leqslant j\leqslant n, a\neq 0$$

称为**第二类初等方阵**. 又

$$P_{jk}(\lambda)=\begin{pmatrix} I^{(j-1)} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \lambda & 0 \\ 0 & 0 & I^{(k-j-1)} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & I^{(n-k)} \end{pmatrix}$$

$$=\sum_{l=1}^{n} E_{ll}+\lambda E_{jk}, \quad 1\leqslant j,k\leqslant n, j\neq k$$

称为**第三类初等方阵**.

显然有

$$P_{jk}^{-1}=P'_{jk}=P_{jk}, \quad P_j(a)^{-1}=P_j(a^{-1}), \quad P_j(a)'=P_j(a),$$
$$P_{jk}(\lambda)^{-1}=P_{jk}(-\lambda), \quad P_{jk}(\lambda)'=P_{kj}(\lambda).$$

所以它们都是非异方阵,且初等方阵之逆方阵仍为初等方阵.

由直接计算不难证明: 给定 $n\times m$ 矩阵 A, 则有

(1) 将 A 的第 j 行(列)与第 k 行(列)互换, 等于矩阵 $P_{jk}^{(n)}A$ ($AP_{jk}^{(m)}$), 这称为对矩阵 A 的行(列)作了**第一类初等变换**;

(2) 将 A 的第 j 行(列)乘以非零常数 a, 等于矩阵 $P_j(a)^{(n)}A$ ($AP_j(a)^{(m)}$), 这称为对矩阵 A 的行(列)作了**第二类初等变换**;

(3) 将 A 的第 k 行(第 j 列)遍乘以 λ, 再加到第 j 行(第 k 列)上, 而第 k 行(第 j 列)不改变, 等于矩阵 $P_{jk}(\lambda)^{(n)}A$ ($AP_{jk}(\lambda)^{(m)}$), 这称为对矩阵 A 的行(列)作了**第三类初等变换**. 三类初等变换统称为**初等变换**.

引理 4.4.1 任给 $n\times m$ 矩阵 A, 则存在一系列初等变换, 它们将 A 变为 $n\times m$ 矩阵

$$A_r=\begin{pmatrix} I^{(r)} & 0^{(r,m-r)} \\ 0^{(n-r,r)} & 0^{(n-r,m-r)} \end{pmatrix}.$$

用式子表出, 即存在 $s+t$ 个初等方阵 $P_1, P_2, \cdots, P_s, Q_1, Q_2, \cdots,$

Q_t,使得

$$P_sP_{s-1}\cdots P_1AQ_1Q_2\cdots Q_t=A_r.$$

证 设$A=0$,则不必证明,设$A\neq0$,则存在$a_{uv}\neq0$. 将第u行和第1行互换(如果$u=1$,不用换),再将第v列和第1列互换(如果$v=1$,也不用换),于是A变成

$$B=\begin{pmatrix}a_{uv} & * \\ * & A_1\end{pmatrix}$$

这就相当于无妨设$a_{11}\neq0$. 这时将A的第1行遍乘以a_{11}^{-1},这就相当于无妨设$a_{11}=1$,即有

$$A=\begin{pmatrix}1 & a_{12} & \cdots & a_{1m}\\ a_{21} & a_{22} & \cdots & a_{2m}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nm}\end{pmatrix}.$$

将A的第1行遍乘以$-a_{j1}$再加到第j行上,$2\leqslant j\leqslant n$,再将第1列遍乘以$-a_{1k}$再加到第k列上,$2\leqslant k\leqslant m$,便得到$n\times m$矩阵

$$B=\begin{pmatrix}1 & 0\\ 0 & B_1\end{pmatrix}.$$

显然对B_1作初等变换等于对B作初等变换. 由归纳法便证明了引理. 证完.

作为推论有

定理 4.4.1 n阶非异方阵为有限个n阶初等方阵的乘积.

证 设A为n阶非异方阵,即有$\det A\neq0$. 由引理4.4.1,所以存在初等方阵$P_1,\cdots,P_s,Q_1,\cdots,Q_t$使得

$$P_s\cdots P_1AQ_1\cdots Q_t=\begin{pmatrix}I^{(r)} & 0^{(r,n-r)}\\ 0^{(n-r,r)} & 0^{(n-r)}\end{pmatrix}.$$

双方取行列式,由于左边不等于零,所以$r=n$,即有

$$A=P_1^{-1}\cdots P_s^{-1}Q_t^{-1}\cdots Q_1^{-1}.$$

由于初等方阵的逆方阵仍为初等方阵. 这证明了定理. 证完.

定义 $n\times m$矩阵A及B称为**相抵的**,如果存在n阶非异方阵P及m阶非异方阵Q,使得

$$B=PAQ.$$

引理 4.4.2 相抵关系为等价关系(见附录2).

证 显然$A=I^{(n)}AI^{(m)}$,即A和A相抵,这证明了**反身性**. 再若$B=PAQ$,由于$\det P\det Q\neq 0$,所以$A=P^{-1}BQ^{-1}$,即若A和B相抵,则B和A相抵,这证明了**对称性**.最后,若$B=PAQ, C=P_1BQ_1$,其中P, P_1为n阶非异方阵,Q, Q_1为m阶非异方阵,于是$C=P_1(PAQ)Q_1=(P_1P)A(QQ_1)$. 这证明了$A$和$C$相抵,即证明了**传递性**. 证完.

定理 4.4.2 任给$n\times m$矩阵A,则唯一存在非负整数r,使得A和标准形$A_r=\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix}$相抵. 所以$A_r$称为**矩阵在相抵下的标准形**,数$r$为**矩阵在相抵下的全系不变量**. (见附录2).

证 由定理4.4.1及引理4.4.1便证明了存在性. 下面证唯一性. 即证明:如果存在n阶非异方阵P, P_1及m阶非异方阵Q, Q_1,使得

$$PAQ=\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix},\quad P_1AQ_1=\begin{pmatrix}I^{(s)} & 0\\ 0 & 0\end{pmatrix}.$$

那末$s=r$,事实上,令

$$A=P^{-1}\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix}Q^{-1}=P_1^{-1}\begin{pmatrix}I^{(s)} & 0\\ 0 & 0\end{pmatrix}Q_1^{-1},$$

于是

$$PP_1^{-1}\begin{pmatrix}I^{(s)} & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix}Q^{-1}Q_1.$$

将n阶非异方阵PP_1^{-1}按前r行,前s列分成四块,将m阶非异方

阵 $Q^{-1}Q_1$ 按前 r 行，前 s 列分成四块．记作

$$PP_1^{-1}=\begin{pmatrix} P_{11}{}^{(r,s)} & P_{12}{}^{(r,n-s)} \\ P_{21}{}^{(n-r,s)} & P_{22}{}^{(n-r,n-s)} \end{pmatrix},$$

$$Q^{-1}Q_1=\begin{pmatrix} Q_{11}{}^{(r,s)} & Q_{12}{}^{(r,m-s)} \\ Q_{21}{}^{(m-r,s)} & Q_{22}{}^{(m-r,m-s)} \end{pmatrix}.$$

于是
$$\begin{pmatrix} P_{11} & 0 \\ P_{21} & 0 \end{pmatrix}=\begin{pmatrix} Q_{11} & Q_{12} \\ 0 & 0 \end{pmatrix}$$

因此 $P_{11}=Q_{11}, Q_{12}=0, P_{21}=0$．即

$$PP_1^{-1}=\begin{pmatrix} P_{11}{}^{(r,s)} & P_{12} \\ 0 & P_{22} \end{pmatrix},\ Q^{-1}Q_1=\begin{pmatrix} Q_{11} & 0 \\ Q_{21} & Q_{22} \end{pmatrix}.$$

由于 PP_1^{-1}，$Q^{-1}Q_1$ 非异，所以将它的行列式按前 r 行（s 列）作 Laplace 展开，便证明了 $r\geqslant s(s\geqslant r)$．所以证明了 $s=r$．定理证完．

注意：定理 4.4.1 也给出了一种计算非异方阵的逆方阵的办法．事实上，由于存在初等方阵 $R_1, R_2, \cdots, R_s$，有

$$R_sR_{s-1}\cdots R_1A=I,\quad A^{-1}=R_sR_{s-1}\cdots R_1,$$

所以考虑 $n\times 2n$ 矩阵 (A, I)，则有

$$R_sR_{s-1}\cdots R_1(A,I)=(I,A^{-1}).$$

即将对 A 施行一系列行的初等变换施行在矩阵 (A, I) 上，使得 A 变为 I，于是单位方阵 I 变为 A^{-1}．

下面引进矩阵的一个重要概念：秩的概念．

定义 $n\times m$ 非零矩阵 A 的所有子式中必有一个阶数最大的非零子式．其阶数称为矩阵 A 的**秩**，记作 $\text{rank}(A)$．$n\times m$ 零矩阵 0 的秩定义为零．

由定义有

引理 4.4.3 $n\times m$ 矩阵 A 的秩有

(i) $0\leqslant \text{rank}(A)\leqslant \min(n, m)$;

(ii) $\text{rank}(\lambda A)=\text{rank}(A)$, $\forall$ 数 $\lambda\neq 0$;

(iii) $\operatorname{rank}(A')=\operatorname{rank}(A)$;

(iV) $\operatorname{rank}(\bar{A})=\operatorname{rank}(A)$.

定义 $n\times m$ 矩阵 A 称为满秩的，如果

$$\operatorname{rank}(A)=\min(m,n).$$

显然有

定理 4.4.3 n 阶方阵 A 非异当且仅当可逆，当且仅当满秩，即对 n 阶方阵，非异、可逆、满秩是互相等价的概念.

下面给出秩的又一个等价定义

定理 4.4.4 $n\times m$ 矩阵 A 的秩为 r 当且仅当在 A 中有一个 r 阶子式不等于零，且所有 $r+1$ 阶子式都等于零，

证 只证明 A 有一个 r 阶子式不等于零，且所有 $r+1$ 阶子式都等于零，则 A 的秩为 r，则 A 的任一 $r+k(k\geqslant1)$ 阶子式都等于零. 实际上，设 B 为 A 的 $r+k,k\geqslant1$ 阶子方阵，将 B 按前 $r+1$ 行作 Laplace 展开. 由条件便证明了 $\det B=0$. 证完.

定理 4.4.5 设 A 为 $n\times m$ 矩阵，B 为 $p\times q$ 矩阵，C 为 $n\times q$ 矩阵. 则有

$$\operatorname{rank}\left(\begin{pmatrix}A & C\\ 0 & B\end{pmatrix}\right)\geqslant\operatorname{rank}\left(\begin{pmatrix}A & 0\\ 0 & B\end{pmatrix}\right)=\operatorname{rank}(A)+\operatorname{rank}(B).$$

证 设 $r=\operatorname{rank}(A)$，$s=\operatorname{rank}(B)$. 于是 A 中存在 r 阶子矩阵 A_0，B 中存在 s 阶子矩阵 B_0，使得 $\det A_0\neq0$，$\det B_0\neq0$. 从而 $\begin{pmatrix}A & 0\\ 0 & B\end{pmatrix}$ 中有 $r+s$ 阶子矩阵 $\begin{pmatrix}A_0 & 0\\ 0 & B_0\end{pmatrix}$ 非异. 由 Laplace 展开定理可知 $\begin{pmatrix}A & C\\ 0 & B\end{pmatrix}$ 中有 $r+s$ 阶子矩阵 $\begin{pmatrix}A_0 & C_0\\ 0 & B_0\end{pmatrix}$ 非异. 这证明了

$$\operatorname{rank}\left(\begin{pmatrix}A & C\\ 0 & B\end{pmatrix}\right)\geqslant\operatorname{rank}(A)+\operatorname{rank}(B),$$

$$\operatorname{rank}\left(\begin{pmatrix}A & 0\\ 0 & B\end{pmatrix}\right)\geqslant\operatorname{rank}(A)+\operatorname{rank}(B).$$

所以问题化为证 rank

$$\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) \leqslant \operatorname{rank}(A) + \operatorname{rank}(B).$$

由定理 4.4.4，考虑$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$的任一 $r+s+1$ 阶子方阵，它必形如

$$X=\begin{pmatrix} A_1 & 0 \\ 0 & B_1 \end{pmatrix}$$

其中 A_1 为 $n_1 \times m_1$ 矩阵，B_1 为 $p_1 \times q_1$ 矩阵. 而

$$n_1+p_1=m_1+q_1=r+s+1.$$

下面证明 $\det X=0$. 用反证法. 设若不然，即 $\det X \neq 0$. 由 Laplace 展开式可证 $n_1=m_1$，$p_1=q_1$，且 $\det X=\det A_1 \det B_1$. 但是由 $n_1+p_1=r+s-1$ 可知 $n_1 \geqslant r+1$ 或者 $p_1 \geqslant s+1$. 当 $n_1 \geqslant r+1$，由 rank $(A)=r$ 可知 $\det A_1=0$；当 $p_1 \geqslant s+1$，由 rank $(B)=s$ 可知 $\det B_1=0$. 总之，证明了 $\det X=0$. 这导出矛盾. 证完.

定理 4.4.6 设 A 为 $n \times m$ 矩阵，B 为 $m \times p$ 矩阵，则有

$$\operatorname{rank}(AB) \leqslant \min(\operatorname{rank}(A), \operatorname{rank}(B)).$$

证 记 rank $(AB)=r$，任取 AB 的 r 阶子式，由 Binet-Cauchy 公式可知，

$$(AB)\begin{pmatrix} i_1 i_2 \cdots i_r \\ j_1 j_2 \cdots j_r \end{pmatrix} = \sum_{1 \leqslant k_1 < \cdots < k_r \leqslant m} A\begin{pmatrix} i_1 i_2 \cdots i_r \\ k_1 k_2 \cdots k_r \end{pmatrix} B\begin{pmatrix} k_2 k_2 \cdots k_r \\ j_1 j_2 \cdots j_r \end{pmatrix}.$$

今 rank $(AB)=r$，所以存在一个 r 阶子式不等于零，设为 $(AB)\begin{pmatrix} i_1 i_2 & \cdots & i_r \\ j_1 j_2 & \cdots & j_r \end{pmatrix} \neq 0$，则存在 $1 \leqslant k_1 < \cdots < k_r \leqslant m$，使得

$$A\begin{pmatrix} i_1 i_2 & \cdots & i_r \\ k_1 \ k_2 & \cdots & k_r \end{pmatrix} B\begin{pmatrix} k_1 k_2 & \cdots & k_r \\ j_1 j_2 & \cdots & j_r \end{pmatrix} \neq 0.$$

于是 $\operatorname{rank}(A) \geqslant r$，$\operatorname{rank}(B) \geqslant r$，这证明了定理. 证完.

定理 4.4.7 设 A 为 $n \times m$ 矩阵，P 为 n 阶非异方阵，

Q 为 m

阶非异方阵，则有

$$\text{rank}(PA)=\text{rank}(AQ)=\text{rank}(A)$$

证 由定理 4.4.6,

$$\begin{aligned}\text{rank}(PA)&\leqslant\min(\text{rank}(P),\text{rank}(A))\\&=\min(n,\text{rank}(A))=\text{rank}(A),\end{aligned}$$

由此可知 $\text{rank}(A)=\text{rank}(P^{-1}(PA))\leqslant\text{rank}(PA)$. 这证明了 $\text{rank}(PA)=\text{rank}(A)$，而

$$\text{rank}(AQ)=\text{rank}(Q'A')=\text{rank}(A')=\text{rank}(A).$$

所以证明了定理. 证完.

定理 4.4.8 设 A 为 $n\times m$ 矩阵，B 为 $m\times p$ 矩阵，则有

$$\text{rank}(A)+\text{rank}(B)\leqslant\text{rank}(AB)+m.$$

证 由定理 4.4.5，有

$$\text{rank}\left(\begin{pmatrix}A & 0\\ I^{(m)} & B\end{pmatrix}\right)\geqslant\text{rank}\left(\begin{pmatrix}A & 0\\ 0 & B\end{pmatrix}\right)=\text{rank}A+\text{rank}B.$$

而
$$\begin{pmatrix}I^{(n)} & -A\\ 0 & I^{(m)}\end{pmatrix}\begin{pmatrix}A & 0\\ I^{(m)} & B\end{pmatrix}=\begin{pmatrix}0 & -AB\\ I^{(m)} & B\end{pmatrix},$$

所以

$$\begin{aligned}\begin{pmatrix}I^{(n)} & -A\\ 0 & I^{(m)}\end{pmatrix}\begin{pmatrix}A & 0\\ I^{(m)} & B\end{pmatrix}\begin{pmatrix}I^{(m)} & -B\\ 0 & I^{(p)}\end{pmatrix}&=\begin{pmatrix}0 & -AB\\ I^{(m)} & B\end{pmatrix}\begin{pmatrix}I^{(m)} & -B\\ 0 & I^{(p)}\end{pmatrix}\\&=\begin{pmatrix}0 & -AB\\ I^{(m)} & 0\end{pmatrix},\end{aligned}$$

又

$$\begin{pmatrix}0 & -AB\\ I^{(m)} & 0\end{pmatrix}\begin{pmatrix}0 & I^{(m)}\\ -I^{(p)} & 0\end{pmatrix}=\begin{pmatrix}AB & 0\\ 0 & I^{(m)}\end{pmatrix}.$$

由定理 4.4.7,

$$\text{rank}A+\text{rank}B\leqslant\text{rank}\left(\begin{pmatrix}A & 0\\ I^{(m)} & B\end{pmatrix}\right)=\text{rank}\left(\begin{pmatrix}AB & 0\\ 0 & I^{(m)}\end{pmatrix}\right).$$

$$=\text{rank}(AB)+m.$$

证完.

这个定理给出了一种重要的矩阵技巧，用来计算有关各种矩阵的秩的不等式.

最后，我们将定理 4.4.1 改写为

定理 4.4.9 $n\times m$ 矩阵 A 在相抵下的全系不变量为 A 的秩 $\text{rank}(A)=r$，而标准形为$\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}$.

习题4.4

1. 试证：n 阶方阵 A 不可逆，即 $\det A=0$ 当且仅当存在 n 阶非零方阵 B,C 使得 $AB=0,CA=0$.

2. 试证：若约定只允许对 $n\times m$ 矩阵 A 的行作初等变换，对 A 的列作第一类初等变换，则可将 A 化为$\begin{pmatrix} I^{(r)} & B \\ 0 & 0 \end{pmatrix}$，其中 $r=\text{rank}(A)$.

3. 试用相抵下的标准形来证明：任取 $n\times m$ 矩阵 A,B，则有
$$\lambda^m\det(\lambda I^{(n)}-AB')=\lambda^n\det(\lambda I^{(m)}-B'A).$$

4. 给定 $n\times m$ 矩阵 A，试求与 A 适合关系 $A'B=B'A$ 的所有 $n\times m$ 矩阵 B.

5. 设 A 和 B 都是 $n\times m$ 矩阵，试证：
$$\text{rank}(A+B)\leqslant\text{rank}(A)+\text{rank}(B),$$
$$\text{rank}(A\bar{A}')=\text{rank}(\bar{A}'A)=\text{rank}(A).$$

6. 设 A 为 $n\times m$ 矩阵，B 为 $m\times p$ 矩阵. 设
$$m=\text{rank}(AB).$$
试证：
$$\text{rank}(A)=\text{rank}(B)=m.$$

7. (i) A 为 n 阶幂等方阵，即 $A^2=A$ 当且仅当
$$\text{rank}(A)+\text{rank}(I-A)=n.$$
(ii) A 为 n 阶对合方阵，即 $A^2=I^{(n)}$ 当且仅当
$$\text{rank}(I+A)+\text{rank}(I-A)=n.$$

8. 试证：$n\times m$ 矩阵 A 的秩等于 1 当且仅当存在 $n\times 1$ 非零矩阵 α 及 $m\times 1$ 非零矩阵 β，使得 $\alpha\beta'=A$.

9. 设 n 阶方阵 A 中任取 s 行构成的 $s\times n$ 矩阵为 B，则有

$$\operatorname{rank}(B)\geqslant\operatorname{rank}(A)+s-n.$$

10. 设 $n\times m$ 矩阵 $A=\begin{pmatrix}B & C\\ D & E\end{pmatrix}$，其中 B 为 r 阶子矩阵，且 $\operatorname{rank}(B)=r$. 试证：$\operatorname{rank}(A)=r$ 当且仅当 $E=DB^{-1}C$.

11. 设 A 为 n 阶方阵，其中至少有 n^2-n+1 个元素为零，试证：$\operatorname{rank}(A)<n$. 这时最大可能的秩是多少？

12. 设自然数 $n\geqslant 2$，且 A 为 n 阶方阵. 记 A^* 为 A 的伴随方阵. 试证：

(i) $\operatorname{rank}(A)=n$ 当且仅当 $\operatorname{rank}(A^*)=n$；

(ii) $\operatorname{rank}(A)=n-1$ 当且仅当 $\operatorname{rank}(A^*)=1$；

(iii) $\operatorname{rank}(A)<n-1$ 当且仅当 $\operatorname{rank}(A^*)=0$.

由此可见，n 阶方阵的伴随方阵的秩只有 $0,1,n$ 之一.

(iv) $(A^*)^*=(\det A)^{n-2}A$.

13. 设 A 和 B 都是 n 阶方阵，则 $n\times 2n$ 矩阵 (A,B) 有

$$\operatorname{rank}((A,B))\leqslant\operatorname{rank}(A)+\operatorname{rank}(B).$$

14. 试证：秩为 r 的 $n\times m$ 矩阵必为 r 个秩为 1 的 $n\times m$ 矩阵之和.

15. 设 A 和 B 都是 n 阶方阵，且有 $\operatorname{rank}(A)+\operatorname{rank}(B)=n$. 试求矩阵方程：$AXB=0$ 的通解.

16. 设 $A_1,A_2,\cdots,A_p$ 为 p 个 n 阶方阵，试证：

$$\sum_{i=1}^{p}\operatorname{rank}(A_i)\leqslant n(p-1)+\operatorname{rank}(A_1A_2\cdots A_p),$$

且证等式可以达到.

17. 设 A,B,C 分别为 $n\times m$，$m\times p$，$m\times q$ 矩阵，且 $AB=0$，$AC=0$，$\operatorname{rank}(A)+\operatorname{rank}(B)=n$. 试证：存在 $p\times q$ 矩阵 D，使得 $C=BD$. 且证唯一存在 D 的必要且充分条件为 $\operatorname{rank}(B)=p$.

18. **Frobenious 不等式**：设 A,B,C 分别为 $n\times m,m\times p,p\times q$ 矩阵. 试证：

$$\operatorname{rank}(AB)+\operatorname{rank}(BC)\leqslant\operatorname{rank}(B)+\operatorname{rank}(ABC).$$

用此式来证：

$$\operatorname{rank}(A)+\operatorname{rank}(B)\leqslant m+\operatorname{rank}(AB),$$

以及当 $\operatorname{rank}(AB)=\operatorname{rank}(B)$ 时有 $\operatorname{rank}(BC)=\operatorname{rank}(ABC)$.

19. 设 A 为 n 阶方阵，且设存在自然数 N，使得

$$\operatorname{rank}(A^N)=\operatorname{rank}(A^{N+1}).$$

试证：

$$\operatorname{rank}(A^N)=\operatorname{rank}(A^{N+1})=\operatorname{rank}(A^{N+2})=\cdots$$

第五章 线性方程组理论

§5.1 非齐次线性方程组

定义 m个未知数 $x_1,x_2,\cdots,x_m$, n个方程的**线性方程组**

$$\begin{cases}a_{11}x_1+\cdots+a_{1m}x_m=a_1,\\ \cdots\cdots\\ a_{n1}x_1+\cdots+a_{nm}x_m=a_n\end{cases}$$

决定了 $n\times m$ 矩阵 A, $n\times 1$ 矩阵 α, $n\times(m+1)$矩阵 $\tilde{A}$, $m\times 1$ 矩阵 x,它们分别定义为

$$A=\begin{pmatrix}a_{11} & \cdots & a_{1m}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nm}\end{pmatrix},\quad \alpha=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix},\quad \tilde{A}=(A,\alpha),\quad x=\begin{pmatrix}x_1\\ \vdots\\ x_m\end{pmatrix}$$

分别称为**系数矩阵**,**常向量**,**增广矩阵**,**变向量**. 这时线性方程组可简洁地写为

$$Ax=\alpha.$$

定义 给定m个未知数, n个方程的线性方程组

$$Ax=\alpha.$$

$m\times 1$ 矩阵

$$x^{(0)}=\begin{pmatrix}x_1^{(0)}\\ \vdots\\ x_m^{(0)}\end{pmatrix}$$

称为线性方程组 $Ax=\alpha$ 的**解**,如果 $Ax^{(0)}=\alpha$.

线性方程组理论就是求出所有解(称为**通解**)的理论,这包含了给出有解的必要且充分条件.

定义 给定m个未知数 $x_1,x_2,\cdots,x_m$, n个方程构成的线性方程组$Ax=\alpha$, 再给定m个未知数$y_1,y_2,\cdots,y_m$, p个方程构成的线性方程组$A_0y=\alpha_0$. 线性方程组$Ax=\alpha$和$A_0y=\alpha_0$称为等价的,如果存在 $1,2,\cdots,m$ 的排列 $i_1i_2\cdots i_m$, 使得令 $y_j=x_{i_j}$, $j=1,2,\cdots,m$后, $Ax=\alpha$的解必为$A_0y=\alpha_0$的解;反之,$A_0y=\alpha_0$的解,必为$Ax=\alpha$的解.

上面定义告诉我们, 如果对系数矩阵A的列作第一类初等变换,变成B,则线性方程组 $By=\alpha$ 和 $Ax=\alpha$ 等价. 一般有

引理 5.1.1 给定线性方程组$Ax=\alpha$. 对系数矩阵A的列作第一类初等变换,对增广矩阵 $\bar{A}=(A,\alpha)$的行作第一、二、三类初等变换. 经过这一系列初等变换后线性方程组变为 $A_0x=\alpha_0$, 则它和原来线性方程组 $Ax=\alpha$ 等价.

证 记m个未知数 $x_1,x_2,\cdots,x_m$, n个方程构成的线性方程组为 $Ax=\alpha$. n个方程可记作 $l_1,l_2,\cdots,l_n$. 对系数矩阵A的列作第一类初等变换,显然使所得方程组与原方程组等价. 增广矩阵$\bar{A}=(A,\alpha)$的行作第一类初等变换,就是将方程 $l_1,l_2,\cdots,l_n$变成 $l_1\cdots,l_k,\cdots,l_j,\cdots,l_n$, $j<k$, 这时通解相同; 对$\bar{A}$的行作第二类初等变换,就是将方程 $l_1,l_2,\cdots,l_n$ 中某个方程 l_i 遍乘非零常数 λ, 这时通解也相同;对$\bar{A}$的行作第三类初等变换,就是将方程 $l_1,l_2,\cdots,l_n$ 变为方程 $l_1,\cdots,l_j+al_k,\cdots,l_k,\cdots,l_n$,其中$a$为常数,这时通解也相同. 证完.

于是有

定理 5.1.1 (Gauss 消去法) 线性方程组 $Ax=\alpha$ 等价于下述形式之线性方程组 $Bx=\beta$,其中

$$B=\begin{pmatrix} B_{11}^{(p)} & B_{12} \\ 0 & 0 \end{pmatrix},\quad B_{11}=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ 0 & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & b_{pp} \end{pmatrix},$$

$b_{11}b_{22}\cdots b_{pp}\neq 0$.

证 只需对增广矩阵 $\tilde{A}=(A,\alpha)$的行作初等变换，对系数矩阵 A的列作第一类初等变换，便不难证明定理. 证完.

定理 5.1.2 (线性方程组的相容性定理) 线性方程组$Ax=\alpha$有解的必要且充分条件为系数矩阵 A 的秩等于增广矩阵 $\tilde{A}=(A,\alpha)$ 的秩. 当有解时，解依赖于 $m-\text{rank}(A)$个独立参数. 特别，当 $m=\text{rank}(A)$时解唯一存在，其中m为独立自变量的个数.

证 由定理 5.1.1，线性方程组 $Ax=\alpha$ 等价于线性方程组 $Bx=\beta$，其中

$$B=\begin{pmatrix} B_{11}^{(p)} & B_{12} \\ 0 & 0 \end{pmatrix},\quad B_{11}=\begin{pmatrix} b_{11} & \cdots & b_{1p} \\ & \ddots & \vdots \\ & & b_{pp} \end{pmatrix},\quad b_{11}b_{22}\cdots b_{pp}\neq 0.$$

于是 $\text{rank}(B)=p$. 由于 B 为由 A 经过行的初等变换及列的第一类初等变换得到，所以秩不改变，即有

$$\text{rank}(A)=p.$$

另一方面，将 β 按前 p 行分块为 $\beta=\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}$，将 x 按前 p 行分块为 $x=\begin{pmatrix}u\\ v\end{pmatrix}$. 于是 $Bx=\beta$ 可写为

$$\begin{pmatrix} B_{11} & B_{12} \\ 0 & 0 \end{pmatrix}\begin{pmatrix}u\\ v\end{pmatrix}=\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}.$$

此即

$$B_{11}u+B_{12}v=\beta_1,\quad \beta_2=0.$$

所以由 $\det B_{11}=b_{11}b_{22}\cdots b_{pp}\neq 0$ 可知

$$u=B_{11}^{-1}\beta_1-B_{11}^{-1}B_{12}v,\quad \beta_2=0.$$

这证明了线性方程组 $Ax=\alpha$ 有解(即线性方程组 $Bx=\beta$ 有解) 当且仅当 $\beta_2=0$，此即增广矩阵

$$(B,\beta)=\begin{pmatrix} B_{11}^{(p)} & B_{12} & \beta_1 \\ 0 & 0 & \beta_2 \end{pmatrix}$$

之秩为 p. 由于增广矩阵 (A,α) 对行作初等变换,再对 A 的列作第一类初等变换变成 (B,β),所以秩不改变. 因此 $Ax=\alpha$ 有解当且仅当 $\text{rank}(A)=p=\text{rank}((A,\alpha))$. 这证明了前一个断言.

当有解时,即 $\beta_2=0$ 时,通解为 $u=B_{11}^{-1}\beta_1-B_{11}^{-1}B_{12}v$, 其中 v 为任意的 $m-p=m-\text{rank}(A)$ 行 1 列矩阵, 即通解依赖于 $m-\text{rank}(A)$ 个独立参数. 证完.

习题 5.1

1. 设 n 阶实方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

适合条件 $a_{ij}\geqslant 0, 1\leqslant i,j\leqslant n$, $\sum_{j=1}^{n} a_{ij}<1$, $1\leqslant i\leqslant n$. 试证: 方阵 $I^{(n)}-A$ 非异,且 $(I-A)^{-1}$ 中每个元素非负.

2. 求下面 n^2 个未知数 $x_{ij}, 1\leqslant i,j\leqslant n$ 的线性方程组的通解, 方程组为

$$a_i a_l x_{jk}-a_j a_k x_{il}=0,\quad i,j,k,l=1,2,\cdots,n.$$

3. 试证:线性方程组 $Ax=\alpha$ 有解的充分条件为矩阵 $\begin{pmatrix} A & \alpha \\ \alpha' & 0 \end{pmatrix}$ 和矩阵 A 的秩相等. 它是不是必要条件,为什么?

4. 设 A 和 B 分别为 $n\times m, n\times p$ 矩阵, X 为由 mp 个独立未知数构成的 $m\times p$ 矩阵. 试证: 矩阵方程 $AX=B$ 有解的必要且充分条件为 $\text{rank}(A)=\text{rank}((A,B))$. 什么时候解唯一存在?

5. 试证: 如果 n 阶方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

有 $2|a_{ii}|>\sum_{j=1}^{n}|a_{ij}|, i=1,2,\cdots,n$，则 $\det A\neq 0$，这时线性方程组 $Ax=\alpha$ 的解唯一存在。

6. 试证：线性方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1m}x_m=a_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+\cdots+a_{nm}x_m=a_n \end{cases}$$

有解的必要且充分条件为对任意 n 个数 $y_1,y_2,\cdots,y_n$，只要 $\sum_{j=1}^{n}y_ja_{ji}=0, i=1,2,\cdots,m$，便有 $\sum_{j=1}^{n}y_jb_j=0$.

§5.2 齐次线性方程组

在这一节考虑 m 个未知数 $x_1,x_2,\cdots,x_m$，n 个方程构成的**齐次线性方程组**

$$Ax=0,$$

这里 A 为 $n\times m$ 矩阵．显然 $x=0$ 为它的解．所以对这类线性方程组要考虑什么时候只有**零解**，什么时候有**非零解**？又通解如何？

定理 5.2.1 (齐次线性方程组解的结构定理) m 个未知数 $x_1,x_2,\cdots,x_m$ 的齐次线性方程组 $Ax=0$ 若有 $\operatorname{rank}(A)=m$，则只有唯一的解 $x=0$；若有 $p=\operatorname{rank}(A)<m$，则有无穷多个解，它依赖于 $m-\operatorname{rank}(A)=m-p$ 个独立参数 $t_{p+1},\cdots,t_m$，使得通解为

$$x=t_{p+1}\beta_{p+1}+\cdots+t_m\beta_m,$$

其中 $\beta_{p+1},\cdots,\beta_m$ 为由 A 确定的 $m-p$ 个 $m\times 1$ 矩阵．特别，$x=\beta_{p+1},\cdots,x=\beta_m$ 都是齐次线性方程组 $Ax=0$ 的解，称为**基础解系**.

证 用定理 5.1.2 于齐次线性方程组 $Ax=0$，即 $\alpha=0$，于是 $\beta=0$．因此 $Ax=0$ 之通解为 $u=-B_{11}^{-1}B_{12}v$，而

$$x=\begin{pmatrix} u \\ v \end{pmatrix}=\begin{pmatrix} -B_{11}^{-1}B_{12}v \\ v \end{pmatrix}.$$

记$(m-p)\times 1$矩阵

$$v=\sum_{j=p+1}^{m} t_j e_j,$$

其中e_j为$(m-p)\times 1$矩阵，其中第$j-p$行元素为1，其余元素为零. 代回去，有

$$x=\sum_{j=p+1}^{m} t_j\begin{pmatrix}-B_{11}^{-1}B_{12}e_j\\ e_j\end{pmatrix}=\sum_{j=p+1}^{m} t_j\beta_j$$

其中 $\beta_j=\begin{pmatrix}-B_{11}^{-1}B_{12}e_j\\ e_j\end{pmatrix}$为$m\times 1$矩阵. 证完.

定理 5.2.2（非齐次线性方程组的解的结构定理） m个未知数的非齐次线性方程组$Ax=\alpha$的解和与它相伴的齐次线性方程组$Ax=0$的解间有关系

(1) $Ax=\alpha$之任两解之差为$Ax=0$之解，

(2) $Ax=\alpha$之解及$Ax=0$之解之和为$Ax=\alpha$之解.

(3) $Ax=\alpha$之通解由$Ax=0$的通解加上$Ax=\alpha$的一个特解构成.

证 (1) 设$m\times 1$矩阵$x^{(1)},x^{(2)}$有$Ax^{(1)}=\alpha$, $Ax^{(2)}=\alpha$, 于是$A(x^{(1)}-x^{(2)})=0$，即$x^{(1)}-x^{(2)}$为$Ax=0$之解.

(2) 设$Ax^{(1)}=\alpha, Ax^{(2)}=0$, 则$A(x^{(1)}+x^{(2)})=\alpha$，即$x^{(1)}+x^{(2)}$为$Ax=\alpha$之解.

(3) 由(1)，(2)可知(3)成立.

证完.

习题 5.2

1. 试证：设$n\times(n+1)$矩阵A的秩为n，则$n+1$个未知数，n个方程构成的线性方程组$Ax=0$的通解为

$$x_k^{(0)}=(-1)^{k-1}tA\begin{pmatrix}12\cdots(k-1)k(k+1)\cdots n\\ 12\cdots(k-1)(k+1)(k+2)\cdots(n+1)\end{pmatrix},\quad k=1,2,\cdots,n+1.$$

§5.3　方阵的特征根

定义　给定n阶方阵A,则关于未知数λ的n次多项式

$$\det(\lambda I^{(n)}-A)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_n$$

称为方阵A的**特征多项式**. 它的n个复根$\lambda_1,\lambda_2,\cdots,\lambda_n$称为方阵$A$的**特征根**. 设$\lambda_0$为$\det(\lambda I-A)$的根,则齐次线性方程组

$$(\lambda_0 I-A)x=0$$

必有非零解,这些解α都称为方阵A的**属于特征根λ_0的特征向量**,即有

$$A\alpha=\lambda_0\alpha,\quad \alpha\neq 0$$

由定义可知

$$a_1=-\mathrm{tr}A=-(\lambda_1+\cdots\lambda_n);$$
$$a_n=(-1)^n\det A=(-1)^n\lambda_1\lambda_2\cdots\lambda_n.$$

且当A为实方阵时,对**实特征根**λ_0,属于λ_0的特征向量α为$n\times 1$非零实矩阵. 一般情形,特征向量为非零复向量.

关于特征多项式有下面重要定理.

定理 5.3.1(Hamilton-Cayley 定理)　n阶方阵A的特征多项式为

$$\det(\lambda I-A)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_n,$$

则有

$$A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nI^{(n)}=0$$

证　由于n阶方阵$\lambda I-A$的伴随方阵$(\lambda I-A)^*$中元素为$\lambda I-A$的代数余子式,所以作为λ的多项式,次数不超过$n-1$. 因此

$$(\lambda I-A)^*=\lambda^{n-1}B_0+\lambda^{n-2}B_1+\cdots+\lambda B_{n-2}+B_{n-1},$$

其中$B_0,B_1,\cdots,B_{n-1}$为n阶常数方阵. 但是$\lambda I-A$的$n-1$阶主子式都是λ的$n-1$次多项式,且首项系数为1,所以B_0的对角元素都是1. 又由 Laplace 展开定理有

$$(\lambda I-A)(\lambda I-A)^*=[\det(\lambda I-A)]\cdot I^{(n)}.$$

所以有

$$(\lambda I-A)(\lambda^{n-1}B_0+\lambda^{n-2}B_1+\cdots+\lambda B_{n-2}+B_{n-1})$$
$$=(\lambda^n+a_1\lambda^{n-1}+\cdots+a_n)I.$$

依次比较 $\lambda^n,\lambda^{n-1},\cdots,\lambda,1$ 之系数，便有

$$B_0=I,$$
$$B_1-AB_0=a_1I,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$B_{n-1}-AB_{n-2}=a_{n-1}I,$$
$$-AB_{n-1}=a_nI.$$

由上往下代入，有

$$B_0=I,$$
$$B_1=a_1I+A,$$
$$B_2=a_2I+a_1A+A^2,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$B_{n-1}=a_{n-1}I+a_{n-2}A+\cdots+a_1A^{n-2}+A^{n-1}.$$

所以

$$a_nI+a_{n-1}A+\cdots+a_1A^{n-1}+A^n=0.$$

这证明了定理. 证完.

定义 给定 n 阶方阵 A，首项系数为 1 的多项式

$$m(\lambda)=\lambda^m+b_1\lambda^{m-1}+\cdots+b_{m-1}\lambda+b_m$$

称为 A 的**极小多项式**，如果

$$m(A)=A^m+b_1A^{m-1}+\cdots+b_{m-1}A+b_mI^{(n)}=0,$$

且是适合这种条件的多项式中次数最小的多项式.

定理 5.3.2 给定 n 阶方阵 A，则有

(1) 唯一存在方阵 A 的极小多项式 $m(\lambda)$;

(2) 设多项式

$$f(\lambda)=c_0\lambda^p+c_1\lambda^{p-1}+\cdots+c_{p-1}\lambda+c_p$$

有

$$f(A)=c_0A^p+c_1A^{p-1}+\cdots+c_{p-1}A+c_pI^{(n)}=0,$$

则 $m(\lambda)\mid f(\lambda)$;

(3) $m(\lambda)\mid\det(\lambda I-A)$,所以极小多项式的根必为特征根;

(4) 反之,方阵 A 的特征根,即特征多项式 $\det(\lambda I-A)$ 的根必为极小多项式的根.

证 由定理 5.3.1 可知存在多项式 $g(\lambda)=\det(\lambda I-A)$,使得 $g(A)=0$. 于是多项式集合

$$\{h(\lambda)\mid h(A)=0\}\neq\{0\}$$

不是空集. 因此在其中存在一个次数最小的,首项系数为 1 的多项式 $m(\lambda)$. 这证明了极小多项式的存在、下面先证(2),再证极小多项式的唯一性. 今任取多项式 $f(\lambda)$,使得 $f(A)=0$. 由 $m(\lambda)\neq 0$,有

$$f(\lambda)=q(\lambda)m(\lambda)+r(\lambda),$$

其中 $r(\lambda)=0$ 或者 $0\leqslant\deg(r(\lambda))<\deg(m(\lambda))$. 但是 $f(A)=0$,此即

$$q(A)m(A)+r(A)=0.$$

已知 $m(A)=0$,所以 $r(A)=0$. 若 $r(\lambda)\neq 0$,由极小多项式定义可知 $\deg(r(\lambda))\geqslant\deg(m(\lambda))$. 这导出矛盾. 所以证明了 $r(\lambda)=0$,此即 $m(\lambda)\mid f(\lambda)$. 这证明了(2). 下面证 (1) 中唯一性. 今若 $m_1(\lambda)$ 的首项系数为 1,且 $m_1(A)=0$. 又 $m_1(\lambda)$ 为次数最小者. 这证明了 $m_1(\lambda)\mid m(\lambda)$, $m(\lambda)\mid m_1(\lambda)$,所以 $m_1(\lambda)=m(\lambda)$. (1) 证完.

由定理 5.3.1 及(2)有 $m(\lambda)\mid\det(\lambda I-A)$,这证明了(3). 余下证(4). 今任取 A 的特征根 λ_0,于是 $\det(\lambda_0I-A)=0$. 所以记

$$x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},$$

则齐次线性方程组$(\lambda_0 I-A)x=0$有非零解α_0，即有

$$A\alpha_0=\lambda_0\alpha_0.$$

记

$$m(\lambda)=\lambda^m+b_1\lambda^{m-1}+\cdots+b_m,$$

则

$$\begin{aligned}0=m(A)\alpha_0&=A^m\alpha_0+b_1A^{m-1}\alpha_0+\cdots+b_{m-1}A\alpha_0+b_m\alpha_0\\&=(\lambda_0^m+b_1\lambda_0^{m-1}+\cdots+b_{m-1}\lambda_0+b_m)\alpha_0=m(\lambda_0)\alpha_0.\end{aligned}$$

由$\alpha_0\neq$，便证明了$m(\lambda_0)=0$，即λ_0为极小多项式$m(\lambda)$的根，这证明了(4)成立. 证完.

最后讨论**圆盘定理**，它给出了特征根的粗略的界.

定理 5.3.3 给定n阶复方阵A，

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix}.$$

记

$$\rho=\max_{1\leqslant i,j\leqslant n}|a_{ij}|,\ \sigma=\frac{1}{2}\max_{1\leqslant i,j\leqslant n}|a_{ij}+\bar{a}_{ji}|,$$

$$\tau=\frac{1}{2}\max_{1\leqslant i,j\leqslant n}|a_{ij}-\bar{a}_{ij}|.$$

则A的特征根λ_0必有

$$|\lambda_0|\leqslant n\rho,\quad |\mathrm{Re}(\lambda_0)|\leqslant n\sigma,\quad |\mathrm{Im}(\lambda_0)|\leqslant n\tau.$$

证 记α_0为n阶方阵A的属于特征根λ_0的特征向量，即有$\alpha_0\neq 0$, $A\alpha_0=\lambda_0\alpha_0$. 于是

$$\bar{\alpha}_0'A\alpha_0=\lambda_0(\bar{\alpha}_0'\alpha_0).$$

记

$$\alpha_0=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}.$$

则有

$$\bar{\alpha}_0'\alpha_0=\sum_{j=1}^{n}|a_j|^2,$$

又

$$|\bar{\alpha}_0'A\alpha_0|=|\sum_{j,k=1}^{n}\bar{a}_j a_{jk}a_k|\leqslant\sum_{j,k=1}^{n}|a_{jk}||a_j||a_k|,$$

所以由 $\rho=\max\limits_{1\leqslant i,j\leqslant n}|a_{ij}|$,有

$$|\bar{\alpha}_0'A\alpha_0|\leqslant\rho\sum_{j,k=1}^{n}|a_j||a_k|=\rho\left(\sum_{j=}^{n}|a_j|\right)^2.$$

由 $\bar{\alpha}_0'A\alpha_0=\lambda_0\bar{\alpha}_0'\alpha_0$,所以证明了

$$|\lambda_0|\sum_{j=1}^{n}|a_j|^2\leqslant\rho\left(\sum_{j=1}^{n}|a_j|\right)^2\leqslant n\rho\sum_{j=1}^{n}|a_j|^2.$$

这证明了 $|\lambda_0|\leqslant n\rho$.

另一方面,有

$$\bar{\alpha}_0'A\alpha_0=\lambda_0(\bar{\alpha}_0'\alpha_0),\qquad \bar{\alpha}_0'\bar{A}'\alpha_0=\bar{\lambda}_0(\bar{\alpha}_0'\alpha_0),$$

于是

$$\bar{\alpha}_0'(A+\bar{A}')\alpha_0=2(\mathrm{Re}(\lambda_0))\bar{\alpha}_0'\alpha_0,$$
$$\bar{\alpha}_0'(A-\bar{A}')\alpha_0=2\sqrt{-1}(\mathrm{Im}(\lambda_0))\bar{\alpha}_0'\alpha_0.$$

和上面一样讨论,便证明了 $|\mathrm{Re}(\lambda_0)|\leqslant n\sigma$, $|\mathrm{Im}(\lambda_0)|\leqslant n\tau$.
证完.

引理 5.3.1 n 阶方阵 A 为 Hermite 方阵当且仅当 $\sqrt{-1}A$ 为斜 Hermite 方阵.

证 由定义直接验证之.

定理 5.3.4 (1) Hermite方阵的特征根都是实数,所以实对称方阵的特征根也都是实数; (2) 斜 Hermite 方阵的非零特征根都是纯虚数,所以实斜对称方阵的非零特征根也都是纯虚数.

证 由引理 5.3.1 可知, 只要证 Hermite 方阵的特征根都是实数就够了. 今

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

是 Hermite 方阵,即 $A=\overline{A}'$,所以

$$a_{ij}=\overline{a_{ji}}.$$

由定理 5.3.3, 任取 A 的特征根 λ_0, 则 $|\mathrm{Im}(\lambda_0)|\leqslant n\tau$, 其中 $\tau=\frac{1}{2}\max\limits_{1\leqslant i,j\leqslant n}|a_{ij}-\overline{a_{ji}}|=0$. 这证明了 $\mathrm{Im}(\lambda_0)=0$. 证完.

定理 5.3.5 (圆盘定理)给定 n 阶方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

记

$$\rho_j=\sum_{k=1}^{n}|a_{jk}|-|a_{jj}|,\qquad \rho_j'=\sum_{k=1}^{n}|a_{kj}|-|a_{jj}|,$$

$$1\leqslant j\leqslant n.$$

在平面上作闭圆盘

$$C_j=\{z\in C\mid |z-a_{jj}|\leqslant\rho_j\},\qquad j=1,2,\cdots,n;$$

$$C_j'=\{z\in C\mid |z-a_{jj}|\leqslant\rho_j'\},\quad j=1,2,\cdots,n.$$

则方阵 A 的特征根必落在并集 $\bigcup\limits_{j=1}^{n}C_j$和并集 $\bigcup\limits_{j=1}^{n}C_j'$ 之交

$$\left(\bigcup_{j=1}^{n}C_j\right)\cap\left(\bigcup_{j=1}^{n}C_j'\right)$$

中.

证 任取 n 阶方阵 A 的特征根 λ_0，由

$$\det(\lambda I-A')=\det(\lambda I-A)$$

可知 λ_0 必为A' 的特征根. 今若能证 $\lambda_0\in\bigcup_{j=1}^{n}C_j$. 则考虑$n$阶方阵$A'$，由 C'_j 之定义可知 $\lambda_0\in\bigcup_{j=1}^{n}C'_j$，这证明了 $\lambda_0\in\left(\bigcup_{j=1}^{n}C_j\right)\cap\left(\bigcup_{j=1}^{n}C'_j\right)$.

下面证 $\lambda_0\in\bigcup_{j=1}^{n}C_j$. 由并集之定义可知，要证存在指标 j，使得 $\lambda_0\in C_j$.

记 α_0 为 n 阶方阵 A 的属于特征根 λ_0 的特征向量，记作

$$\alpha_0=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$$

由定义，$A\alpha_0=\lambda_0\alpha_0$，所以

$$\sum_{j=1}^{n}a_{ij}a_j=\lambda_0a_i,\quad i=1,2,\cdots n,$$

此即

$$\sum_{j\neq i}a_{ij}a_j=(\lambda_0-a_{ii})a_i,\quad i=1,2,\cdots,n.$$

但是 $\alpha_0\neq0$，所以在$|a_1|$，$|a_2|$，$\cdots$，$|a_n|$中存在一个最大数，记作$|a_{i_0}|$，于是

$$\max_{1\leqslant i\leqslant n}|a_i|=|a_{i_0}|>0$$

因此有

$$|\lambda_0-a_{i_0i_0}|\,|a_{i_0}|=\left|\sum_{j\neq i_0}a_{i_0j}a_j\right|$$

$$\leqslant |a_{i_0}| \sum_{j \neq i_0} |a_{i_0 j}| = |a_{i_0}| \rho_{i_0}.$$

由 $|a_{i_0}| \neq 0$，便证明了

$$|\lambda_0 - a_{i_0 i_0}| \leqslant \rho_{i_0},$$

即 $\lambda_0 \in C_{i_0} \subset \bigcup_{j=1}^{n} C_j$. 证完.

习题 5.3

1. 设 n 阶方阵 A 的 n 个特征根为 $\lambda_1, \lambda_2, \cdots, \lambda_n$. 则 A 非异当且仅当 $\lambda_1 \lambda_2 \cdots \lambda_n \neq 0$，这时 A^{-1} 的 n 个特征根为 $\lambda_1^{-1}, \lambda_2^{-1}, \cdots, \lambda_n^{-1}$.

2. 设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 n 阶方阵 A 的 n 个特征根，$f(\lambda)$ 为未知数 λ 的多项式. 试证：n 阶方阵 $f(A)$ 恰好以 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$ 为 n 个特征根. 试用 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 表出 $\det f(A)$.

3. 设 λ_0 为 n 阶方阵 A 的特征根. 记

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

在形式和号

$$\sum_{j=1}^{n} \frac{\sum_{k \neq j} |a_{jk}|^2}{|\lambda_0 - a_{jj}|^2}$$

中除去 $\lambda_0 = a_{jj}$ 及 $\sum_{k \neq j} |a_{jk}|^2 = 0$ 同时成立的项. 试证：它大于或等于 1.

4. 设 α 为 $n \times 1$ 矩阵. 试求 n 阶方阵 $\alpha\alpha'$ 和 $\alpha\bar{\alpha}'$ 的特征根.

5. "记 $f(\lambda) = \det(\lambda I - A)$. 由 $f(A) = \det(A \cdot I - A) = \det O = 0$ 可知定理 5.3.1 成立."这个证明对吗？为什么.

6. 试证：n 阶方阵 A 的特征根 λ_0 有

$$|\lambda_0| \leqslant \min\left(\max_{1 \leqslant l \leqslant n} \left(\sum_{k=1}^{n} |a_{kl}| \right), \max_{1 \leqslant l \leqslant n} \sum_{k=1}^{n} |a_{lk}| \right).$$

7. 试证：幂等方阵的特征根只有 0 和 1；幂零方阵的特征根都是零. 且 n 阶方阵 A 幂零当且仅当 $\mathrm{tr}(A^k) = 0$，$k = 1, 2, \cdots$. 特别，若 n 阶方阵 A 和

$C=AB-BA$ 可交换，则 $C^n=0$.

8. 记 $n_1,\cdots,n_s$ 阶方阵 $A_1,A_2,\cdots,A_s$ 的极小多项式分别为 $m_1(\lambda)$, $m_2(\lambda),\cdots,m_s(\lambda)$. 试证：准对角方阵

$$\mathrm{diag}(A_1,\cdots,A_s)$$

的极小多项式为 $m_1(\lambda),\cdots,m_s(\lambda)$ 的最小公倍式 $[m_1,\cdots,m_s]$.

9. 试证：n 阶可逆方阵 A 的逆方阵 A^{-1} 为 A 的多项式.

10. 设 K 为实斜对称方阵，则对任一正数 a，有

$$\det(aI+K)>0.$$

11. 设 n 阶方阵 $A=(a_{ij})$ 有 $\sum_{j=1}^{n}a_{ij}=1, i=1,2,\cdots,n$，则 A 必有特征根 1. 设 A 有特征根 $\lambda_0\neq 1$. 记 α_0 为 A 的属于特征根 λ_0 的特征向量，则 $n\times 1$ 矩阵 α_0 的元素之和为零.

第六章　线 性 空 间

§6.1　n 维线性空间

下面引进抽象的线性空间

定义　集合 Ω 称为**复(实、有理)线性空间**，如果在 Ω 中有两种代数运算：**加法**"+"和"**纯量乘积**，使得任取 α、$\beta\in\Omega$，则有 $\alpha+\beta\in\Omega$；又任取 $\alpha\in\Omega$，数 a 为复(实、有理)数，有 $a\alpha\in\Omega$. 且适合下面一系列条件：

(i) 加法交换律

$$\alpha+\beta=\beta+\alpha,\quad \forall\alpha,\beta\in\Omega;$$

(ii) 加法结合律

$$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma),\quad \forall\alpha,\beta,\gamma\in\Omega;$$

(iii) 存在**零向量** $0\in\Omega$，使得

$$\alpha+0=0+\alpha=\alpha,\quad \forall\alpha\in\Omega;$$

(iv) 每个向量 $\alpha\in\Omega$，存在**负向量** $-\alpha$，使得

$$\alpha+(-\alpha)=(-\alpha)+\alpha=0,$$

于是可以引进减法如下：

$$\alpha-\beta=\alpha+(-\beta),\quad \forall\alpha,\beta\in\Omega;$$

(v) $(a+b)\alpha=a\alpha+b\alpha,\quad \forall\alpha\in\Omega$，$a,b$ 为复(实、有理)数；

(vi) $a(\alpha+\beta)=a\alpha+a\beta,\quad \forall\alpha,\beta\in\Omega$，$a$ 为复(实、有理)数；

(vii) $a(b\alpha)=(ab)\alpha,\quad \forall\alpha\in\Omega$，$a,b$ 为复(实、有理)数；

(viii) $1\cdot\alpha=\alpha,\quad \forall\alpha\in\Omega.$

这时 Ω 中元素称为**向量**，复(实、有理)数称为**纯量**.

例 1 普通空间情形.由解析几何课程可知,所有自由向量构成一个线性空间,它是实线性空间. 在空间中取定坐标系,熟知每个自由向量有坐标(a_1,a_2,a_3). 于是可作一个 3×1 实矩阵

$$\alpha=\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}.$$

因此所有自由向量构成的线性空间,可以自然地诱导出它们的坐标表达,即由所有 3×1 实矩阵构成的线性空间,且这时加法和纯量乘积就是矩阵的加法和纯量乘积.

例 2 记 V_n 为所有 $n\times1$ 矩阵构成的集合,它们在矩阵的加法及纯量乘积下构成线性空间.

例 3 记 $M(n\times m)$为所有 $n\times m$ 矩阵构成的集合,它们在矩阵的加法及纯量乘积下构成线性空间.

例 4 记 $P[x]$为所有多项式构成的集合,它们在多项式的加法及乘以常数为纯量乘积下构成线性空间.

例 5 记 $P_n[x]$为所有次数不超过 n 的多项式构成的集合,它们在多项式的加法及乘以常数为纯量乘积下构成线性空间.

例 6 记$C([0,1])$为区间$[0,1]$上所有连续函数构成的集合,它们在函数相加及乘以常数为纯量乘积下构成线性空间.

在进一步展开线性空间理论以前,先从定义出发给出关于运算的一些最基本的性质.

引理 6.1.1 设 Ω 为线性空间,则有

(1) 任意有限个向量相加,可以不计先后和次序;

(2) 加法消去律成立,即由 $\alpha+\beta=\alpha+\gamma$, $\forall\alpha,\beta,\gamma\in\Omega$ 可推出 $\beta=\gamma$;

(3) Ω 中有且只有一个零向量. 且对每个向量α,有且只有一个负向量$-\alpha$;

(4) 记 a 为数，$\alpha\in\Omega$，则 $a\alpha=0$ 当且仅当 $a=0$ 或者$\alpha=0$. 因此若 a 和 b,c 为数，$a\neq 0,\alpha\in\Omega$，则由 $ab\alpha=ac\alpha$ 可推出 $b\alpha=c\alpha$;

(5) 记 a 为数，$\alpha\in\Omega$，则有

$$(-a)\alpha=a(-\alpha)=-(a\alpha).$$

证 由加法结合律成立，可知有限个向量相加可以不计先后，由加法交换律成立，可知可以不计次序．这证明了(1)．下面证加法消去律成立，今由 $\alpha+\beta=\alpha+\gamma$，由于等量加等量仍为等量，所以 $(-\alpha)+(\alpha+\beta)=(-\alpha)+(\alpha+\gamma)$．由加法结合律有 $((-\alpha)+\alpha)+\beta=((-\alpha)+\alpha)+\gamma$，由负向量的定义可知 $(-\alpha)+\alpha=0$，故有 $0+\beta=0+\gamma$，由零向量的定义可知 $\beta=\gamma$，这证明了(2).

证(3)．由线性空间定义可知零向量 0 存在．若另有一个零向量 0_1，由定义，则有 $0_1+\alpha=\alpha+0_1=\alpha$, $\forall\alpha\in\Omega$．特别，由 $0\in\Omega$，所以有 $0_1+0=0+0_1=0$．但是 0 为零向量，故 $0+0_1=0_1+0=0_1$代入便证明了 $0_1=0$．这证明了零向量的唯一性．现在任给 $\alpha\in\Omega$，由线性空间定义可知负向量$-\alpha$ 存在．若另有一个负向量 β，由定义，则有 $\alpha+\beta=\beta+\alpha=0$．但是 $\alpha+(-\alpha)=(-\alpha)+\alpha=0$，所以有 $\alpha+\beta=0=\alpha+(-\alpha)$．由加法消去律成立，故证明了 $\beta=-\alpha$．这证明了负向量的唯一性.

证(4)．今 $a\alpha=0$，若 $a\neq 0$，则 a^{-1} 存在，于是

$$a^{-1}0=a^{-1}(a\alpha)=(a^{-1}a)\alpha=1\cdot\alpha=\alpha.$$

但是 $0+0=0$，于是 $a^{-1}0=a^{-1}(0+0)=a^{-1}0+a^{-1}0$．由加法消去律成立，所以$a^{-1}0=0$，这证明了$\alpha=0$．即当$a\alpha=0$ 时有$a=0$或者$\alpha=0$．反之，上面已证 $a0=0$，余下要证 $0\alpha=0$．今 $0\alpha=(0+0)\alpha=0\alpha+0\alpha$，由加法消去律成立，便证明了 $0\alpha=0$．总之当 $a=0$ 或 $\alpha=0$，便有 $a\alpha=0$．所以证明了 $a\alpha=0$ 当且仅当 $a=0$ 或 $\alpha=0$．今若 $a\neq 0$，$ab\alpha=ac\alpha$，于是 $ab\alpha-ac\alpha=0$，因此 $a(b\alpha-c\alpha)=0$．由 $b\alpha-c\alpha\in\Omega$ 及 $a\neq 0$，所以有 $b\alpha-c\alpha=0$，即 $b\alpha=c\alpha$．这证明了(4).

证(5)，今 $a\alpha+(-a)\alpha=(a+(-a))\alpha=0\alpha=0$，所以 $(-a)\alpha=-(a\alpha)$. 又 $a\alpha+a(-\alpha)=a(\alpha+(-\alpha))=a0=0$，所以 $a(-\alpha)=-(a\alpha)$. 这证明了(5). 证完.

上面证明中，每步都是根据定义条件，或定义推出的性质推得所要的结论，这就是所谓的"说话要有根据". 从定义出发，通过逻辑推理，派生出整个数学模式是研究代数的重要方法.

为了展开线性空间理论，最重要的概念为下面三个概念：线性相关、线性无关和线性组合.

定义 线性空间 $\mathfrak{L}$ 中 m 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 称为**线性相关的**，如果存在 m 个不全为零的数 $a_1,a_2,\cdots,a_m$，使得

$$a_1\alpha_1+a_2\alpha_2+\cdots+a_m\alpha_m=\sum_{j=1}^{m}a_j\alpha_j=0.$$

否则 m 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 称为**线性无关的**.

定义 线性空间 $\mathfrak{L}$ 中向量 α 称为 m 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的**线性组合**，如果存在 m 个数 $a_1,a_2,\cdots,a_m$，使得

$$\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_m\alpha_m=\sum_{j=1}^{m}a_j\alpha_j$$

由定义可知，包含零向量的任意有限个向量必线性相关，特别零向量必线性相关. 证明如下：给定 m 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$. 设 $\alpha_i=0$. 我们取 $a_i=1,a_j=0,j\neq i$，$1\leqslant j\leqslant m$，则由引理 6.1.1 之(4)，有 $a_1\alpha_1+\cdots+a_i\alpha_i+\cdots+a_m\alpha_m=0$. 由于 $a_1,\cdots,a_i,\cdots,a_m$ 不全为零，这证明了 $\alpha_1,\cdots,\alpha_i,\cdots,\alpha_m$ 线性相关. 由定义又可知一个非零向量必线性无关. 用反证法来证明之. 今若 $\alpha\neq0$ 线性相关，则存在数 $a\neq0$，有 $a\alpha=0$. 这和引理 6.1.1 之(4)矛盾. 所以证明了：当 $\alpha\neq0$，则 α 线性无关，一般有

定理 6.1.1 线性空间 $\mathfrak{L}$ 中给定 m 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$，如果其中有部分向量线性相关，则全体线性相关；如果全体线性无关，

则它们的任何一部分向量也必线性无关.

证 显然用前一断言及反证法，可证后一断言成立. 下面证前一断言成立.无妨设 $\alpha_1,\cdots,\alpha_p$ 线性相关,于是存在纯量 $a_1,\cdots,a_p$ 不全为零,且有 $\sum_{j=1}^{p} a_j\alpha_j=0$. 今取 $a_{p+1}=\cdots=a_m=0$,于是 $\sum_{j=1}^{m} a_j\alpha_j=0$,且 $a_1,\cdots,a_p,a_{p+1},\cdots,a_m$ 也不全为零,这证明了 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关. 证完.

定理 6.1.2 线性空间 $\mathfrak{L}$ 中 m 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性无关当且仅当若 m 个数 $a_1,a_2,\cdots,a_m$ 有

$$\sum_{j=1}^{m} a_j\alpha_j=0,$$

则有 $a_1=\cdots=a_m=0$.

证 由于 $\alpha_1,\cdots,\alpha_m$ 线性相关的定义为存在 $a_1,a_2,\cdots,a_m$ 不全为零,且 $\sum_{j=1}^{m} a_j\alpha_j=0$, 所以线性无关当且仅当不存在 $a_1,a_2,\cdots,a_m$ 不全为零,且 $\sum_{j=1}^{m} a_j\alpha_j=0$, 此即由 $\sum_{j=1}^{m} a_j\alpha_j=0$ 可推出 $a_1=\cdots=a_m=0$. 证完.

定理 6.1.3 线性空间 $\mathfrak{L}$ 中 m 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关当且仅当其中必有一个向量为其余向量的线性组合.

证 今若 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关，于是存在不全为零的一组数 $a_1,a_2,\cdots,a_m$, 使得 $\sum_{j=1}^{m} a_j\alpha_j=0$. 今若 $a_i\neq0$,则由

$\alpha_i=-a_i^{-1}(a_1\alpha_1+\cdots+a_{i-1}\alpha_{i-1}+a_{i+1}\alpha_{i+1}+\cdots+a_m\alpha_m)$ 可知 α_i 为 $\alpha_1,\cdots,\alpha_{i-1},\alpha_{i+1},\cdots,\alpha_m$ 的线性组合. 反之,不妨设 α_1 为 $\alpha_2,\cdots,\alpha_m$ 的线性组合,即有

$$\alpha_1=b_2\alpha_2+\cdots+b_m\alpha_m,$$

其中 $b_2,\cdots,b_m$ 为数，则有 $(-1)\alpha_1+b_2\alpha_2+\cdots+b_m\alpha_m=0$. 今 -1, $b_2,\cdots,b_m$ 不全为零，于是 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性相关. 证完.

利用上面性质，我们来引进维数概念.

若 $\mathfrak{L}$ 只由一个零向量组成，则 $\mathfrak{L}$ 称为 0 维线性空间，用 0 来表示. 若 $\mathfrak{L}\neq 0$，则其中存在非零向量 α_1. 如果 $\mathfrak{L}$ 中任一向量 β 和 α_1 线性相关，即存在纯量 a,b 不全为零，且 $a\alpha_1+b\beta=0$，自然 $b\neq 0$，因为若 $b=0$，则 $a\alpha_1=0$，但是 $\alpha\neq 0$，所以 $a=0$，这和 a,b 不全为零矛盾. 因此证明了 $\beta=\left(-\frac{a}{b}\right)\alpha_1$，即为 α_1 的纯量倍，所以

$$\mathfrak{L}\subset\{a\alpha_1 \quad | \quad \forall \text{ 数 } a\}.$$

但是由于 $\mathfrak{L}$ 为线性空间，由 $\alpha_1\in\mathfrak{L}$ 可知 $a\alpha_1\in\mathfrak{L}$ 对一切数 a 成立. 这证明了 $\mathfrak{L}=\{a\alpha_1 \mid \forall \text{ 数 } a\}$. 这时 $\mathfrak{L}$ 称为一维线性空间，α_1 称为 $\mathfrak{L}$ 的基. 如果 $\mathfrak{L}$ 中并非任意向量都和 α_1 线性相关，那末存在向量 α_2 和 α_1 线性无关. 于是又出现两种情形：一为 $\mathfrak{L}$ 中任意向量都和 α_1,α_2 线性相关；另一为 $\mathfrak{L}$ 中存在向量 α_3 和 α_1,α_2 线性无关，这样依次讨论下去，用归纳法的思想，便出现了两种情形：

(1) 在线性空间 $\mathfrak{L}$ 中找不到有限个向量，它们线性无关，且 $\mathfrak{L}$ 中任一向量和它们线性相关，这种线性空间称为**无限维线性空间**. 关于无限维线性空间的理论，属于泛函分析这个学科，今后我们不再讨论.

(2) 在线性空间 $\mathfrak{L}$ 中找到 n 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$，它们线性无关，且 $\mathfrak{L}$ 中任一向量 β 和 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性相关. 这类线性空间称为**有限维线性空间**. 它们有下面性质：

定理 6.1.4 设线性空间 $\mathfrak{L}$ 中存在 n 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$，它们线性无关，且 $\mathfrak{L}$ 中任一向量 β 和 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性相关. 则 β 可唯一地表为 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的线性组合

$$\beta=b_1\alpha_1+b_2\alpha_2+\cdots+b_n\alpha_n.$$

且 $\mathfrak{L}$ 中任意 $n+1$ 个向量必定线性相关. 非负整数 n 称为线性空间 $\mathfrak{L}$ 的**维数**,记作 $\dim\mathfrak{L}$. $\alpha_1, \alpha_2, \cdots, \alpha_n$ 称为 $\mathfrak{L}$ 的一组**基**. β 关于 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的组合系数排成 $n\times 1$ 矩阵

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in V_n$$

称为向量 β 关于基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的**坐标**, b_j 称为 β 在 α_j 方向的**分量**,或第 j 个(坐标)分量, $j=1,2,\cdots,n$.

证 由条件, β 和 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性相关. 于是存在 $n+1$ 个不全为零的数 $b, a_1, a_2, \cdots, a_n$, 使得

$$b\beta + a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n = 0.$$

我们来证 $b\neq 0$. 设若不然,即 $b=0$, 于是 $a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n = 0$, 由定理 6.1.2, $a_1 = a_2 = \cdots = a_n = 0$, 这和 $b, a_1, a_2, \cdots, a_n$ 不全为零矛盾,所以证明了 $b\neq 0$, 因此

$$\beta = \left(-\frac{a_1}{b}\right)\alpha_1 + \left(-\frac{a_2}{b}\right)\alpha_2 + \cdots + \left(-\frac{a_n}{b}\right)\alpha_n.$$

取 $b_j = -\dfrac{a_j}{b}$, $j=1,2,\cdots,n$. 便证明了 $\beta = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_n\alpha_n$.

再证分解唯一性,设 $\beta = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_n\alpha_n = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$, 因此有 $(b_1-c_1)\alpha_1 + (b_2-c_2)\alpha_2 + \cdots + (b_n-c_n)\alpha_n = 0$. 由定理 6.1.2, 所以 $b_1=c_1$, $b_2=c_2$, $\cdots$, $b_n=c_n$. 这证明了分解唯一性.

最后,任取 $n+1$ 个向量 $\beta_1, \beta_2, \cdots, \beta_{n+1}$, 于是有

$$\beta_j = \sum_{k=1}^{n} a_{jk}\alpha_k, \qquad j=1,2,\cdots,n+1.$$

下面证 $\beta_1, \beta_2, \cdots, \beta_{n+1}$ 线性相关,即证存在 $n+1$ 个不全为零的数 $x_1^{(0)}, x_2^{(0)}, \cdots, x_{n+1}^{(0)}$ 使得 $\sum\limits_{j=1}^{n+1} x_j^{(0)}\beta_j = 0$. 事实上, 引进 $n+1$ 个独立未知

数 $x_1, x_2, \cdots, x_{n+1}$，则向量方程$\sum_{j=1}^{n+1} x_j \beta_j = 0$ 可写为

$$0 = \sum_{j=1}^{n+1} x_j \beta_j = \sum_{j=1}^{n+1} x_j \left(\sum_{k=1}^{n} a_{jk} \alpha_k \right) = \sum_{k=1}^{n} \left(\sum_{j=1}^{n+1} a_{jk} x_j \right) \alpha_k.$$

由于 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关，所以向量方程变为齐次线性方程组

$$\sum_{j=1}^{n+1} a_{jk} x_j = 0, \quad k = 1, 2, \cdots, n.$$

它有 $n+1$ 个未知数，n 个方程，它的系数矩阵为 $n \times (n+1)$ 矩阵，所以秩 $r \leqslant n$，由 § 5.2，它的通解依赖于 $n+1-r \geqslant 1$ 个独立参数。因此证明了它有非零解，即 $\beta_1, \beta_2, \cdots, \beta_{n+1}$ 线性相关．定理证完．

注意 维数概念与基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的选取无关，只与线性空间 $\mathfrak{L}$ 有关．事实上，若另有 m 个向量 $\delta_1, \delta_2, \cdots, \delta_m$ 线性无关，而任一向量与 $\delta_1, \delta_2, \cdots, \delta_m$ 线性相关，由于 $\mathfrak{L}$ 中任意 $n+1$ 个向量线性相关，所以 $m \leqslant n$．和定理 6.1.3 一样可证，$\mathfrak{L}$ 中任意 $m+1$ 个向量线性相关，所以 $n \leqslant m$，因此证明了 $m=n$，所以维数和基的选取无关．

习题 6.1

1. 试证：空间解析几何中引进的自由向量全体构成三维实线性空间．又在其中任取 m 个自由向量 $\alpha_1, \alpha_2, \cdots, \alpha_m$，使得 α_{j-1} 的终点为 α_j 之起点，$j = 2, 3, \cdots, m$，又 α_m 之终点为 α_1 的起点．试证：$\sum_{j=1}^{m} \alpha_j = 0$.

2. 中国象棋中的马从某一点起跳，经过若干步跳回原起跳点．试证：它共跳了偶数步．

3. 设 $r \geqslant 2$. 在线性空间 $\mathfrak{L}$ 中取线性相关的向量组 $\alpha_1, \alpha_2, \cdots, \alpha_r$，再取一向量 β. 试证：存在 r 个不全为零的数 $a_1, a_2, \cdots, a_r$，使得向量组

$$\alpha_1 + a_1 \beta, \cdots, \alpha_r + a_r \beta$$

线性相关．

4. 若向量

$$\alpha_j=(a_{j1},\cdots,a_{jk})\in V_k,\quad 1\leqslant j\leqslant r$$

线性无关，则

$$\beta_j=(a_{j1},\cdots,a_{jk},a_{j,k+1},\cdots,a_{jn})\in V_n,\quad 1\leqslant j\leqslant r$$

也线性无关

5. n 阶方阵 A 的 n 个行及 n 个列，作为 $n\times 1$ 矩阵，是线性空间 V_n 中向量，分别称为 A 的行向量及列向量，试证：A 非异当且仅当它的 n 个行向量线性无关，当且仅当它的 n 个列向量也线性无关.

6. 试证：n 阶方阵的属于不同特征根的特征向量线性无关.

7. 记 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 n 维线性空间 $\mathfrak{L}$ 的一组基．如果 $\beta\in\mathfrak{L}$，且 β 为 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 中任意 $n-1$ 个向量的线性组合，试证：$\beta=0$.

8. 记 $\beta_1,\beta_2,\cdots,\beta_m$ 为 n 维线性空间 $\mathfrak{L}$ 中线性相关向量组，但是其中任取 $m-1$ 个向量必线性无关．设 m 个纯量 $b_1,b_2,\cdots,b_m$ 有 $\sum\limits_{j=1}^{m}b_j\beta_j=0$，则 $b_1b_2\cdots b_m\neq 0$ 或者 $b_1=\cdots=b_m=0$. 在前一情形如果另有纯量 $c_1,c_2,\cdots,c_m$，使得 $\sum\limits_{j=1}^{m}c_j\beta_j=0$，则有 $b_1:c_1=b_2:c_2=\cdots=b_m:c_m$.

§6.2 基及基变换

定理 6.2.1 设 $\mathfrak{L}$ 为 n 维线性空间，则任取 r 个线性无关的向量 $\alpha_1,\cdots,\alpha_r$，必有 $r\leqslant n$，且在 $\mathfrak{L}$ 中必存在 $n-r$ 个向量 $\alpha_{r+1},\cdots,\alpha_n$，使得 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$ 为 $\mathfrak{L}$ 的一组基.

证 由定理 6.1.4 可知 $r\leqslant n$. 当 $r=n$ 不用证了．当 $r<n$. 如果不存在向量 α_{r+1}，使得 $\alpha_1,\cdots,\alpha_r,\ \alpha_{r+1}$ 线性无关，则对 $\mathfrak{L}$ 中任一向量 β，那末 $\alpha_1,\cdots,\alpha_r,\ \beta$ 线性相关．和定理 6.1.4 一样可证 $\mathfrak{L}$ 中任意 $r+1$ 个向量必线性相关，但是 $\dim\mathfrak{L}=n$，即 $\mathfrak{L}$ 中存在 n 个向量线性无关，这证明了 $n<r+1$．它和 $r<n$ 矛盾．所以在 $\mathfrak{L}$ 中存在 α_{r+1} 使得 $\alpha_1,\cdots,\ \alpha_r,\alpha_{r+1}$ 线性无关．依次讨论下去便证明了定理．证完.

定理 6.2.2 在 n 维线性空间 Ω 中取定两组基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 和 $\{\beta_1, \beta_2, \cdots, \beta_n\}$，则基变换公式：

$$\beta_i = \sum_{j=1}^{n} a_{ji}\alpha_j, \quad i=1,2,\cdots,n$$

决定的 n 阶方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

非异. 反之，任取 n 阶非异方阵 $A=(a_{ij})$，构造 n 个向量

$$\beta_i = \sum_{j=1}^{n} a_{ji}\alpha_j, \quad 1 \leqslant i \leqslant n$$

则 $\beta_1, \beta_2, \cdots, \beta_n$ 为 Ω 的基. 因此，Ω 中的基和 n 阶非异方阵在基变换公式下一一对应.

证 令 $\{\alpha_1, \cdots, \alpha_n\}$，$\{\beta_1, \cdots, \beta_n\}$ 为 Ω 中两组基. 于是有

$$\beta_i = \sum_{j=1}^{n} a_{ji}\alpha_j, \qquad \alpha_i = \sum_{j=1}^{n} b_{ji}\beta_j, \quad 1 \leqslant i \leqslant n.$$

因此

$$\alpha_i = \sum_{j=1}^{n} b_{ji}\beta_j = \sum_{j=1}^{n}\sum_{k=1}^{n} b_{ji}a_{kj}\alpha_k = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{kj}b_{ji}\right)\alpha_k,$$
$$1 \leqslant i \leqslant n.$$

由于 $\alpha_1, \cdots, \alpha_n$ 线性无关，所以有

$$\sum_{j=1}^{n} a_{kj}b_{ji} = \delta_{ik}, \quad i,k=1,2,\cdots,n.$$

记

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \qquad B=\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}.$$

这证明了 $AB=I^{(n)}$，所以 $\det A\neq 0$.

反之，若 $\det A\neq 0, \beta_i=\sum_{j=1}^{n}a_{ji}\alpha_j,\ i=1,\ 2,\ \cdots,\ n$. 我们来证 $\beta_1,\beta_2,\cdots,\beta_n$ 为 Ω 中一组基，即证 $\beta_1,\beta_2,\cdots,\beta_n$ 线性无关. 事实上，若线性相关，则存在一组不全为零的常数 $x_1^{(0)},\cdots,x_n^{(0)}$，使得 $\sum_{j=1}^{n}x_j^{(0)}\beta_j=0$. 于是

$$0=\sum_{j=1}^{n}x_j^{(0)}\beta_j=\sum_{j=1}^{n}\sum_{k=1}^{n}x_j^{(0)}a_{kj}\alpha_k=\sum_{k=1}^{n}\left(\sum_{j=1}^{n}a_{kj}x_j^{(0)}\right)\alpha_k.$$

由于 $\alpha_1,\cdots,\alpha_n$ 线性无关，所以有

$$\sum_{j=1}^{n}a_{kj}x_j^{(0)}=0,\quad k=1,2,\cdots,n.$$

由于 $\det A\neq 0$ 及 Cramer 法则，便证明了 $x_1^{(0)}=\cdots=x_n^{(0)}=0$，这和 $x_1^{(0)},\cdots,x_n^{(0)}$不全为零矛盾. 所以证明了 $\beta_1,\cdots,\beta_n$ 线性无关. 证完.

定理 6.2.3 在 n 维线性空间 Ω 中取两组基 $\{\alpha_1,\ \cdots,\ \alpha_n\}$及 $\{\beta_1,\cdots,\beta_n\}$. 在 Ω 中取定向量 α，设 α 关于基$\{\alpha_1,\cdots,\alpha_n\}$的坐标为 x，关于基$\{\beta_1,\cdots,\beta_n\}$的坐标为 y，则有如下**坐标变换公式**

$$y=A^{-1}x,$$

其中 $A=(a_{ij})$为基变换公式 $\beta_i=\sum_{j=1}^{n}a_{ji}\alpha_j, 1\leqslant i\leqslant n$ 对应的 n阶非异方阵.

证 今记

$$x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\qquad y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},$$

于是

$$\alpha=\sum_{j=1}^{n} x_j\alpha_j=\sum_{l=1}^{n} y_l\beta_l.$$

而

$$\sum_{l=1}^{n} y_l\beta_l=\sum_{l=1}^{n}\sum_{j=1}^{n} y_l a_{jl}\alpha_j=\sum_{j=1}^{n}\left(\sum_{l=1}^{n} y_l a_{jl}\right)\alpha_j,$$

由于 $\alpha_1,\cdots,\alpha_n$ 为基，所以有

$$x_j=\sum_{l=1}^{n} a_{jl}y_l.$$

即有 $x=Ay$，所以 $y=A^{-1}x$. 证完.

习题 6.2

1. 给定线性空间 Ω 中 n 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$，它们适合条件

(1) $\alpha=\sum_{j=1}^{n} a_j\alpha_j, \forall\alpha\in\Omega$，即 α 为 $\alpha_1,\cdots,\alpha_n$ 的线性组合;

(2) 存在向量 $\alpha_0=\sum_{j=1}^{n} b_j\alpha_j$，且它关于 $\alpha_1,\cdots,\alpha_n$ 之线性组合系数唯一.

试证: $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 Ω 的一组基.

2. 在 n 维线性空间 Ω 中取定两组基 $\{\alpha_1,\cdots,\alpha_n\}$ 及 $\{\beta_1,\cdots,\beta_n\}$. 试证: 存在 $1,2,\cdots,n$ 的排列 $i_1i_2\cdots i_n$，使得

$$\{\alpha_1,\cdots,\alpha_{j-1},\beta_{i_j},\alpha_{j+1},\cdots,\alpha_n\}, j=1,2,\cdots,n$$

都是 Ω 的基.

§6.3 同　构

定义 给定线性空间 Ω_1 和 Ω_2，它们的纯量同时限制在复(实、有理)数的范围. 记 σ 为 Ω_1 到 Ω_2 内的单值映射. σ 称为 Ω_1 到 Ω_2 内的**线性映射**，如果它有

(1) $\sigma(\alpha+\beta)=\sigma(\alpha)+\alpha(\beta)$, $\forall\alpha,\beta\in\Omega_1$;

(2) $\sigma(a\alpha)=a\sigma(\alpha)$, $\quad\forall\alpha\in\mathfrak{L}$, 数 a.

当 $\mathfrak{L}_2=\mathfrak{L}_1$ 时，线性映射改称为**线性变换**.

在下一章，我们专门讨论线性映射与线性变换. 现在只考虑下面特殊的线性映射.

定义 线性空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 的纯量约定在相同的复（实、有理）数的范围. 线性空间 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 上的线性映射称为**同构映射**，如果它还是到上的一一映射. 这时 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 称为**同构**，记作 $\mathfrak{L}_1\simeq\mathfrak{L}_2$.

引理 6.3.1 复(实、有理）数上的线性空间之间的同构关系为等价关系(见附录 2)

证 作用恒等映射可知线性空间 $\mathfrak{L}$ 和 $\mathfrak{L}$ 同构，这证明了反身性. 今若 $\sigma:\mathfrak{L}_1\to\mathfrak{L}_2$ 为同构映射. 显然 σ^{-1}：$\mathfrak{L}_2\to\mathfrak{L}_1$ 仍为同构映射，这证明了当 $\mathfrak{L}_1\simeq\mathfrak{L}_2$，则有 $\mathfrak{L}_2\simeq\mathfrak{L}_1$，即证明了对称性. 最后，若 $\sigma:\mathfrak{L}_1\to\mathfrak{L}_2$、$\tau:\mathfrak{L}_2\to\mathfrak{L}_3$ 都是同构映射，则 $\tau\circ\sigma:\mathfrak{L}_1\to\mathfrak{L}_3$ 仍为同构映射，这证明了：由 $\mathfrak{L}_1\simeq\mathfrak{L}_2$, $\mathfrak{L}_2\simeq\mathfrak{L}_3$ 可知 $\mathfrak{L}_1\simeq\mathfrak{L}_3$，即传递性成立. 所以同构关系为等价关系. 证完.

定理 6.3.1 在 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，对 $\mathfrak{L}$ 中任一向量 $\beta=\sum\limits_{j=1}^{n}b_j\alpha_j$，于是建立了 $\mathfrak{L}$ 到 V_n 上的对应关系

$$\sigma:\ \beta=\sum_{j=1}^{n}b_j\alpha_j\to\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}\in V_n.$$

则 σ 为 $\mathfrak{L}$ 到 V_n 上的同构映射，即 $\mathfrak{L}\simeq V_n$. 因此任一 n 维线性空间必同构于 V_n. 所以线性空间在同构关系下的全系不变量为它的维数.

证 由于 $\alpha_1,\cdots,\alpha_n$ 为 $\mathfrak{L}$ 的基，所以 β 的坐标表达唯一. 这证明了 σ 为 $\mathfrak{L}$ 到 V_n 上的一一对应. 容易直接验证 σ 为 $\mathfrak{L}$ 到 V_n 的同

构映射．证完．

由此定理可知，在同构意义下，只要考虑线性空间V_n就行了．

§ 6.4 子 空 间

定义 给定n维线性空间$\mathfrak{L}$，设$\mathfrak{S}$为$\mathfrak{L}$中一个子集合．如果$\mathfrak{S}\neq\{0\}$，且$\mathfrak{S}$中有r个向量$\alpha_1,\alpha_2,\cdots,\alpha_r$线性无关，而$\mathfrak{S}$中任一向量$\beta$和$\alpha_1,\cdots,\alpha_r$线性相关，则称$\alpha_1,\alpha_2,\cdots,\alpha_r$为$\mathfrak{S}$中**极大线性无关部分组**．

引理 6.4.1 设$\mathfrak{S}$为n维线性空间$\mathfrak{L}$中子集合．如果$\mathfrak{S}\neq\{0\}$，$\alpha_1,\cdots,\alpha_r$为$\mathfrak{S}$中极大线性无关部分组，则$\mathfrak{S}$中任一向量为$\alpha_1,\cdots,\alpha_r$之线性组合，且$1\leqslant r\leqslant n$．

证 由定理 6.1.4 可知$1\leqslant r\leqslant n$．任取$\beta\in\mathfrak{S}$，由极大线性无关组之定义可知β，$\alpha_1,\cdots,\alpha_r$线性相关．于是存在不全为零的数$b,a_1,\cdots,a_r$，使得$b\beta+a_1\alpha_1+\cdots+a_r\alpha_r=0$．由$\alpha_1,\cdots,\alpha_r$线性无关可知$b\neq0$，因此$\beta$为$\alpha_1,\cdots,\alpha_r$的线性组合．证完．

定义 给定n维线性空间$\mathfrak{L}$，$\mathfrak{L}$中子集合$\mathfrak{L}_1$称为**子空间**，如果有

(1) $\alpha+\beta\in\mathfrak{L}_1$，　$\forall\alpha,\beta\in\mathfrak{L}_1$，

(2) $a\alpha\in\mathfrak{L}_1$，　　$\forall\alpha\in\mathfrak{L}_1$，数$a$．

引理 6.4.2 n维线性空间$\mathfrak{L}$的子空间$\mathfrak{L}_1$为线性空间，当$\mathfrak{L}_1\neq\{0\}$，则$\mathfrak{L}_1$的极大线性无关部分组$\alpha_1,\cdots,\alpha_r$为线性空间$\mathfrak{L}_1$的基．所以$\dim\mathfrak{L}_1=r\leqslant n=\dim\mathfrak{L}$．又在$\mathfrak{L}_1$中任取基$\beta_1,\cdots,\beta_r$，则存在$\mathfrak{L}$中向量$\beta_{r+1},\cdots,\beta_n$使得$\beta_1,\cdots,\beta_r,\beta_{r+1},\cdots,\beta_n$为$\mathfrak{L}$的基．

证 由子空间及线性空间的定义，由极大线性无关部分组的定义以及定理 6.1.4 便证明了引理．证完．

定义 给定n维线性空间$\mathfrak{L}$．记$\mathfrak{S}$为$\mathfrak{L}$中子集合．在$\mathfrak{S}$中任取有限个向量作各种可能的线性组合，它们全体构成的集合记

作$[\mathfrak{S}]$,则$[\mathfrak{S}]$称为由 $\mathfrak{S}$ **线性生成**的子集合.

引理 6.4.3 设 $\mathfrak{S}$ 为 n 维线性空间 $\mathfrak{L}$ 的子集合,则由 $\mathfrak{S}$ 线性生成的子集合 $[\mathfrak{S}]$ 为 $\mathfrak{L}$ 的子空间. 当 $\mathfrak{S}=\{0\}$, 则$[\mathfrak{S}]=\{0\}$; 当 $\mathfrak{S}\neq\{0\}$,则 $\mathfrak{S}$ 的极大线性无关部分组为子空间$[\mathfrak{S}]$的基.

证 由线性生成的定义及子空间的定义可知$[\mathfrak{S}]$为 $\mathfrak{L}$ 的子空间,显然当 $\mathfrak{S}=\{0\}$,则$[\mathfrak{S}]=\{0\}$. 设 $\mathfrak{S}\neq\{0\}$,设 $\alpha_1,\cdots,\alpha_r$ 为 $\mathfrak{S}$ 的极大线性无关部分组. 于是 $\mathfrak{S}$ 中任一元为 $\alpha_1,\cdots,\alpha_r$ 的线性组合,因此 $\mathfrak{S}$ 中有限个元素的线性组合仍为 $\alpha_1,\cdots,\alpha_r$ 的线性组合,所以$[\mathfrak{S}]$中任一元为 $\alpha_1,\cdots,\alpha_r$ 的线性组合, 这证明了 $\alpha_1,\cdots,\alpha_r$ 为$[\mathfrak{S}]$的一组基. 证完.

注意引理 6.4.1 和引理 6.4.3 并没有回答是否存在极大线性无关部分组. 这个回答是肯定的,事实上,我们有

引理 6.4.4 设 $\mathfrak{S}\neq\{0\}$ 为 n 维线性空间 $\mathfrak{L}$ 中子集合, 则在 $\mathfrak{S}$ 中存在极大线性无关部分组; 且 $\mathfrak{S}$ 中任两极大线性无关部分组中向量之个数相同,记作 r,则 $1\leqslant r\leqslant n$,且

$$\dim([\mathfrak{S}])=r.$$

证 今 $\mathfrak{S}\neq\{0\}$, 所以在 $\mathfrak{S}$ 中存在非零向量 α_1, 它线性无关. 如果 $\mathfrak{S}$ 中任一向量与 α_1 线性相关,则$\{\alpha_1\}$为极大线性无关部分组. 否则, 在 $\mathfrak{S}$ 中存在向量 α_2 和 α_1 线性无关, 注意到 $\dim\mathfrak{L}=n$,所以 $\mathfrak{L}$ 中任意 $n+1$ 个向量线性相关, 因此 $\mathfrak{S}$ 中任意 $n+1$ 个向量线性相关,所以我们从 α_1 出发,依次选取下去,最后可证存在 r 个向量 $\alpha_1,\cdots,\alpha_r$ 线性无关, 且 $\mathfrak{S}$ 中任一向量和 $\alpha_1,\cdots,\alpha_r$ 线性相关, 这证明了 $\alpha_1,\cdots,\alpha_r$ 为 $\mathfrak{S}$ 的极大线性无关部分组,自然 $1\leqslant r\leqslant n$.

由引理 6.4.3, $\mathfrak{S}$ 中任一组极大线性无关部分组为由 $\mathfrak{S}$ 线性生成的子空间$[\mathfrak{S}]$的基, 这证明了 $\mathfrak{S}$ 的任一组极大线性无关部分组中向量之个数都等于 $\dim([\mathfrak{S}])$. 证完.

定义 在 n 维线性空间 $\mathfrak{L}$ 中取定向量组$\{\alpha_1,\cdots,\alpha_r\}$和$\{\beta_1,\cdots,$

$\beta_s\}$. 如果每个 $\beta_j, j=1,\cdots,s$ 都是 $\alpha_1, \cdots, \alpha_r$ 的线性组合，则称 $\{\beta_1,\cdots,\beta_s\}$ 由 $\{\alpha_1,\cdots,\alpha_r\}$ **线性表出**.

定理 6.4.1（替换定理） 在 n 维线性空间 Ω 中取定 m 个线性无关的向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$. 任给向量组 $\{\beta_1,\cdots,\beta_s\}$. 如果 $\{\alpha_1,\cdots,\alpha_m\}$ 由 $\{\beta_1,\cdots,\beta_s\}$ 线性表出，则 $m\leqslant s$. 且在 $\{\beta_1,\cdots,\beta_s\}$ 中存在子集合 $\{\beta_{i_1},\cdots,\beta_{i_m}\}$，$1\leqslant i_1<\cdots<i_m\leqslant s$，使得在子集合 $\{\beta_1,\cdots,\beta_s\}$ 中除去向量 $\beta_{i_1},\cdots,\beta_{i_m}$，用 $\alpha_1,\cdots,\alpha_m$ 的替换，这样得到的新的向量组和 $\{\beta_1,\cdots,\beta_s\}$ 可以互相线性表出.

证 由条件

$$\alpha_i=\sum_{j=1}^{s}a_{ji}\beta_j,\quad i=1,2,\cdots,m$$

于是有 $s\times m$ 矩阵

$$A=\begin{pmatrix}a_{11} & \cdots & a_{1m}\\ \vdots & & \vdots\\ a_{s1} & \cdots & a_{sm}\end{pmatrix}.$$

引进 m 个独立未知数 $x_1,\cdots,x_m$. 于是由 $\sum_{i=1}^{m}x_i\alpha_i=0$ 有 $x_1=0,\cdots,x_m=0$. 今

$$\sum_{i=1}^{m}x_i\alpha_i=\sum_{i=1}^{m}x_i\sum_{j=1}^{s}a_{ji}\beta_j=\sum_{j=1}^{s}\left(\sum_{i=1}^{m}x_ia_{ji}\right)\beta_j$$

考虑齐次线性方程组

$$\sum_{i=1}^{m}a_{ji}x_i=0,\quad j=1,2,\cdots,s.$$

如果它有非零解 $x_1=x_1^{(0)},\cdots,x_m=x_m^{(0)}$，那末 $x_1^{(0)},\cdots,x_m^{(0)}$ 不全为零，且 $\sum_{i=1}^{m}x_i\alpha_i=0$. 这导出矛盾，所以证明了这个齐次线性方程组只有零解. 此即 $m-\mathrm{rank}(A)=0$，因此证明了 $s\times m$ 矩阵

A 的秩为 m，于是 $m \leqslant s$，且在 A 中存在一个 m 阶子式不等于零。为讨论方便起见，无妨假设

$$\det A_0 = \det \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix} \neq 0.$$

记 $B = A_0^{-1} = (b_{ij})$，则有：对 $j = 1, 2, \cdots, m$，

$$\begin{aligned} \beta_j &= \sum_{k=1}^{m} \delta_{jk} \beta_k = \sum_{k=1}^{m} \left(\sum_{l=1}^{m} a_{kl} b_{lj} \right) \beta_k \\ &= \sum_{l=1}^{m} b_{lj} \left(\sum_{k=1}^{m} a_{kl} \beta_k \right) = \sum_{l=1}^{m} b_{lj} \left(\alpha_l - \sum_{k=m+1}^{s} a_{kl} \beta_k \right) \end{aligned}$$

这证明了 $\{\beta_1, \cdots, \beta_m, \beta_{m+1}, \cdots, \beta_s\}$ 可由 $\{\alpha_1, \cdots, \alpha_m, \beta_{m+1}, \cdots, \beta_s\}$ 线性表出. 证完.

于是有

定理 6.4.2 线性空间 Ω 中两个线性无关向量组 $\{\alpha_1, \cdots, \alpha_m\}$，$\{\beta_1, \cdots, \beta_s\}$ 如果可以互相线性表出，则有 $m = s$.

证 由定理 6.4.1，有 $m \leqslant s$，$s \leqslant m$. 证完.

定理 6.4.3 给定 $n \times m$ 矩阵

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix},$$

则 n 个行向量

$$\alpha_i = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{im} \end{pmatrix} \in V_m, \quad i = 1, 2, \cdots, n$$

的极大线性无关部分组中向量之个数等于 rank(A)，且 m 个列向量

$$\beta_j=\begin{pmatrix}a_{1j}\\ \vdots\\ a_{nj}\end{pmatrix}\in V_n,\quad j=1,2,\cdots,m$$

的极大线性无关部分组中向量之个数也等于 rank(A).

证 由于 rank(A)=rank(A'). 而A的列向量为 A' 的行向量，所以只要证明前一断言就够了. 为讨论方便起见，无妨设 $A\binom{1\ 2\ \cdots\ r}{1\ 2\ \cdots\ r}\neq 0$, $r=\text{rank}(A)$. 下面是证 $\alpha_1,\cdots,\alpha_r$ 线性无关. 事实上，如果线性相关，则存在 r 个数 $a_1,\cdots,a_r$，它们不全为零，且 $\sum_{i=1}^{r}a_i\alpha_i=0$. 所以 $\sum_{i=1}^{r}a_ia_{ij}=0$, $1\leqslant j\leqslant m$. 这证明了齐次线性方程组 $\sum_{i=1}^{r}a_{ij}x_i=0$, $1\leqslant j\leqslant m$ 有非零解. 而 $\sum_{i=1}^{r}a_{ij}x_i=0$, $j=1,\cdots,r$ 的系数矩阵的行列式为 $A\binom{1\ 2\ \cdots\ r}{1\ 2\ \cdots\ r}\neq 0$. 由 Cramer 法则可知它只有零解. 这和 $a_1,\cdots,a_r$ 不全为零矛盾，所以 $\alpha_1,\cdots,\alpha_r$ 线性无关. 显然为了证明定理，只要证明$\{\alpha_1,\cdots,\alpha_r\}$为$\{\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n\}$ 的极大线性无关部分组就够了，即只要证明 α_j 与 $\alpha_1,\cdots,\alpha_r$ 线性相关，其中$r+1\leqslant j\leqslant n$. 引进 $r+1$ 个独立未知数 $x_j,x_1,\cdots,x_r$，若

$$\sum_{i=1}^{r}x_i\alpha_i+x_j\alpha_j=0,$$

则有

$$\sum_{i=1}^{r}x_ia_{ik}+x_ja_{jk}=0,\quad k=1,2,\cdots,m$$

它的系数矩阵为

$$\begin{pmatrix}a_{11}&a_{21}&\cdots&a_{r1}&a_{j1}\\ \vdots&\vdots&&\vdots&\vdots\\ a_{1m}&a_{2m}&\cdots&a_{rm}&a_{jm}\end{pmatrix},$$

秩为 r，未知数个数为 $r+1$. 所以解依赖于 $(r+1)-r=1$ 个参数. 这证明了它一定有非零解，所以 $\alpha_1,\cdots,\alpha_r,\alpha_j$ 线性相关，$r+1\leqslant j\leqslant n$. 证完.

推论 设 $n\times m$ 矩阵 A 的秩为 r，则 A 的所有行向量线性生成了 V_m 中 r 维子空间，所有列向量线性生成了 V_n 中 r 维子空间.

定义 记 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 为 n 维线性空间 $\mathfrak{L}$ 中两个子空间，则子集合 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 之交 $\mathfrak{L}_1\cap\mathfrak{L}_2$ 仍为子空间，称为 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 的**交**.

由定义可知，任意多个子空间之交仍为子空间.

定义 记 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 为 n 维线性空间 $\mathfrak{L}$ 的两个子空间，则 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 之**和**为子集合

$$\mathfrak{L}_1+\mathfrak{L}_2=\{\alpha+\beta\mid\forall\alpha\in\mathfrak{L}_1,\ \beta\in\mathfrak{L}_2\}.$$

由定义可知子空间之和仍为子空间，所以有限多个子空间之和仍为子空间，且关于子空间之交及和还有性质

(1) $\mathfrak{L}_1\cap\mathfrak{L}_2=\mathfrak{L}_2\cap\mathfrak{L}_1$;

(2) $(\mathfrak{L}_1\cap\mathfrak{L}_2)\cap\mathfrak{L}_3=\mathfrak{L}_1\cap(\mathfrak{L}_2\cap\mathfrak{L}_3)$;

(3) $\mathfrak{L}_1+\mathfrak{L}_2=\mathfrak{L}_2+\mathfrak{L}_1$;

(4) $(\mathfrak{L}_1+\mathfrak{L}_2)+\mathfrak{L}_3=\mathfrak{L}_1+(\mathfrak{L}_2+\mathfrak{L}_3)$;

(5) 记 0 为由零向量构成之子空间，则有

$$\mathfrak{L}_1+0=0+\mathfrak{L}_1=\mathfrak{L}_1.$$

关于维数，还有

定理 6.4.4 设 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 为 n 维线性空间 $\mathfrak{L}$ 的子空间，则有

$$\dim\mathfrak{L}_1+\dim\mathfrak{L}_2=\dim(\mathfrak{L}_1\cap\mathfrak{L}_2)+\dim(\mathfrak{L}_1+\mathfrak{L}_2).$$

证 已知 $\mathfrak{L}_1\cap\mathfrak{L}_2$ 为 $\mathfrak{L}_1$ 及 $\mathfrak{L}_2$ 的子空间. 在 $\mathfrak{L}_1\cap\mathfrak{L}_2$ 中取定基 $\alpha_1,\cdots,\alpha_r$，在 $\mathfrak{L}_1$ 中补足 $\beta_1\cdots,\beta_s$，使 $\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s$ 为 $\mathfrak{L}_1$ 的基. 在 $\mathfrak{L}_2$ 中补足 $\gamma_1,\cdots,\gamma_t$，使 $\alpha_1,\cdots,\alpha_r,\gamma_1,\cdots,\gamma_t$ 为 $\mathfrak{L}_2$ 的基. 于是 $\dim\mathfrak{L}_1+\dim\mathfrak{L}_2-\dim(\mathfrak{L}_1\cap\mathfrak{L}_2)=r+s+t$. 所以为了证明定理，

只要证 $\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s,\gamma_1,\cdots,\gamma_t$ 为 $\mathfrak{L}_1+\mathfrak{L}_2$ 之基就够了．由 $\mathfrak{L}_1+\mathfrak{L}_2$ 之定义可知，$\mathfrak{L}_1+\mathfrak{L}_2$ 中向量都是 $\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s,\gamma_1,\cdots,\gamma_t$ 之线性组合，于是问题化为证明它们线性无关．今若

$$\sum_{i=1}^{r} a_i\alpha_i+\sum_{j=1}^{s} b_j\beta_j+\sum_{k=1}^{t} c_k\gamma_k=0,$$

则 $\sum_{k=1}^{t} c_k\gamma_k=-\sum_{i=1}^{r} a_i\alpha_i-\sum_{j=1}^{s} b_j\beta_j\in\mathfrak{L}_1\cap\mathfrak{L}_2$．由于 $\alpha_1,\cdots,\alpha_r,\gamma_1,\cdots,\gamma_t$ 线性无关，便证明了 $c_1,\cdots,c_t=0$，于是 $\sum_{i=1}^{r} a_i\alpha_i+\sum_{j=1}^{s} b_j\beta_j=0$．由 $\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s$ 线性无关，所以 $a_1=\cdots=a_r=0,b_1=\cdots=b_s=0$．这证明了 $\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s,\gamma_1,\cdots,\gamma_t$ 线性无关，即为 $\mathfrak{L}_1+\mathfrak{L}_2$ 之基．证完．

定义 n 维线性空间 $\mathfrak{L}$ 中子空间 $\mathfrak{L}_1$ 及 $\mathfrak{L}_2$ 之和 $\mathfrak{L}_1+\mathfrak{L}_2$ 称为**直接和**，如果 $\mathfrak{L}_1\cap\mathfrak{L}_2=0$．这时记作 $\mathfrak{L}_1\oplus\mathfrak{L}_2$．

推广这个定义，有

定义 n 维线性空间 $\mathfrak{L}$ 中子空间 $\mathfrak{L}_1,\mathfrak{L}_2,\cdots,\mathfrak{L}_t$ 之和

$$\mathfrak{L}_0=\mathfrak{L}_1+\mathfrak{L}_2+\cdots+\mathfrak{L}_t=\sum_{j=1}^{t}\mathfrak{L}_j$$

称为**直接和**，如果

$$\dim\mathfrak{L}_0=\dim\mathfrak{L}_1+\dim\mathfrak{L}_2+\cdots+\dim\mathfrak{L}_t$$

这时记直接和为

$$\mathfrak{L}_0=\mathfrak{L}_1\oplus\mathfrak{L}_2\oplus\cdots\oplus\mathfrak{L}_t=\sum_{j=1}^{t}{}'\mathfrak{L}_j.$$

定理 6.4.5 设 $\mathfrak{L}_1,\mathfrak{L}_2,\cdots,\mathfrak{L}_t$ 为 n 维线性空间 $\mathfrak{L}$ 的子空间．关于 $\mathfrak{L}_1,\mathfrak{L}_2,\cdots,\mathfrak{L}_t$ 之和 $\mathfrak{L}_0=\sum_{j=1}^{t}\mathfrak{L}_j$，下面四性质互相等价(即直接和有四种互相等价的定义)

(1) $\dim \mathfrak{L}_0 = \sum_{j=1}^{t} \dim \mathfrak{L}_j$;

(2) 设 $\alpha_{j1}, \cdots, \alpha_{jn_j}$ 为 $\mathfrak{L}_j$ 的基, $j=1,2,\cdots,t$, 则

$$\alpha_{11}, \cdots, \alpha_{1n_1}, \cdots, \alpha_{t1}, \cdots, \alpha_{tn_t}$$

为 $\mathfrak{L}_0$ 的基;

(3) 若 $\sum_{j=1}^{t} \alpha_j = 0$, 其中 $\alpha_1 \in \mathfrak{L}_1, \cdots, \alpha_t \in \mathfrak{L}_t$, 则有 $\alpha_j = 0$, $j=1, 2, \cdots, t$;

(4) 任取 $\alpha \in \mathfrak{L}_0$, 则 α 可唯一分解为 $\alpha = \sum_{j=1}^{t} \alpha_j$, 其中 $\alpha_1 \in \mathfrak{L}_1, \cdots, \alpha_t \in \mathfrak{L}_t$. 即若 $\alpha = \sum_{j=1}^{t} \beta_j$, 其中 $\beta_1 \in \mathfrak{L}_1, \cdots, \beta_t \in \mathfrak{L}_t$, 便有 $\beta_j = \alpha_j$, $1 \leqslant j \leqslant t$.

证 先证(1)和(2)等价. 今若 $\dim \mathfrak{L}_0 = \sum_{j=1}^{t} \dim \mathfrak{L}_j$. 由于和的定义, $\mathfrak{L}_0$ 中向量为 $\alpha_{11}, \cdots, \alpha_{1n_1}, \cdots, \alpha_{t1}, \cdots, \alpha_{tn_t}$ 的线性组合. 于是 $\alpha_{11}, \cdots, \alpha_{1n_1}, \cdots, \alpha_{t1}, \cdots, \alpha_{tn_t}$ 中极大线性无关部分组为 $\mathfrak{L}_0$ 之基, 由 $\dim \mathfrak{L}_0 = \sum_{j=1}^{t} \dim \mathfrak{L}_t$, 所以极大线性无关部分组之个数 $\dim \mathfrak{L}_0 = \sum_{j=1}^{t} n_j$. 这证明了 $\alpha_{11}, \cdots, \alpha_{1n_1}, \cdots, \alpha_{t1}, \cdots, \alpha_{tn_t}$ 线性无关, 即为 $\mathfrak{L}_0$ 之基. 反之, 若 $\alpha_{11}, \cdots, \alpha_{1n_1}, \cdots, \alpha_{t1}, \cdots, \alpha_{tn_t}$ 为 $\mathfrak{L}_0$ 之基, 于是 $\dim \mathfrak{L}_0 = \sum_{j=1}^{t} n_j = \sum_{j=1}^{t} \dim \mathfrak{L}_j$. 证完.

下面证由(2)可推出(3), 由(3)可推出(4), 由(4)可推出(2). 于是(2), (3), (4)互相等价.

先由(2)推(3). 今 $\alpha_j \in \mathfrak{L}_j$, 所以 $\alpha_j = \sum_{k=1}^{n_j} a_{jk} \alpha_{jk}$, $1 \leqslant j \leqslant t$, 而

$0=\sum_{j=1}^{t}\alpha_j=\sum_{j=1}^{t}\sum_{k=1}^{n_j}a_{jk}\alpha_{jk}$. 由于(2)成立，即$a_{11},\cdots,\alpha_{1n_1},\cdots,\alpha_{t1},\cdots,\alpha_{tn_t}$为基，所以$a_{jk}=0$，$1\leqslant k\leqslant n_j$，$1\leqslant j\leqslant t$. 这证明了$\alpha_j=0$，$1\leqslant j\leqslant t$.

再由(3)推(4). 若$\alpha_j,\beta_j\in\mathfrak{L}_j$，$1\leqslant j\leqslant t$，且$\alpha=\sum_{j=1}^{t}\alpha_j=\sum_{j=1}^{t}\beta_j$. 于是$\sum_{j=1}^{t}(\alpha_j-\beta_j)=0$，其中$\gamma_j=\alpha_j-\beta_j\in\mathfrak{L}_j$，$1\leqslant j\leqslant t$. 由$\sum_{j=1}^{t}\gamma_j=0$及(3)可知$\gamma_j=0$，此即$\beta_j=\alpha_j$，$1\leqslant j\leqslant t$. 这证明了分解的唯一性.

最后由(4)推(2)，已知任取$\alpha\in\mathfrak{L}_0$，则有$\alpha=\sum_{j=1}^{t}\alpha_j$，且分解唯一，其中$\alpha_j\in\mathfrak{L}_j$，$1\leqslant j\leqslant t$. 在$\mathfrak{L}_j$中取基$\alpha_{j1},\cdots,\alpha_{jn_j}$，$j=1,2,\cdots,t$. 显然$\alpha_{11},\cdots,\alpha_{1n_1},\cdots,\alpha_{t1},\cdots,\alpha_{tn_t}$线性生成$\mathfrak{L}_0$. 问题化为证明它们是$\mathfrak{L}_0$的一组基，今若有数$a_{jk}$，$1\leqslant k\leqslant n_j$，$1\leqslant j\leqslant t$，有$\sum_{j=1}^{t}\sum_{k=1}^{n_j}a_{jk}\alpha_{jk}=0$. 由于$\sum_{k=1}^{n_j}a_{jk}\alpha_{jk}=\alpha_j\in\mathfrak{L}_j$，于是$0=\alpha_1+\cdots+\alpha_t$. 由分解唯一性，有$\alpha_j=0$，$1\leqslant j\leqslant t$. 但是$\alpha_{j1},\cdots,\alpha_{jn_j}$为$\mathfrak{L}_j$的基，所以由$0=\alpha_j=\sum_{k=1}^{n_i}a_{jk}\alpha_k$有$a_{jk}=0$，$1\leqslant k\leqslant n_j$，$1\leqslant j\leqslant t$. 这证明了$\alpha_{11},\cdots,\alpha_{1n_1},\cdots,\alpha_{t1},\cdots,\alpha_{tn_t}$为$\mathfrak{L}_0$的一组基，证完.

推论 当$t=2$时，上面关于直接和的两种定义互相等价. 即$\dim\mathfrak{L}_1+\dim\mathfrak{L}_2=\dim(\mathfrak{L}_1+\mathfrak{L}_2)$当且仅当$\mathfrak{L}_1\cap\mathfrak{L}_2=0$.

证 这是维数公式$\dim\mathfrak{L}_1+\dim\mathfrak{L}_2=\dim\mathfrak{L}_1\cap\mathfrak{L}_2+\dim(\mathfrak{L}_1+\mathfrak{L}_2)$的直接推论. 证完.

习题 6.4

1. 试判断 n 维线性空间 V_n 中下列子集合是否构成子空间?如果是，试确定其维数并找出一组基; 如果不是，试写出它线性生成的子空间，并确定维数和找出一组基.

(i) $\{(a_1, \cdots, a_r, a_{r+1}, \cdots, a_n) \mid a_1 + \cdots + a_r = 0\}$;

(ii) n 个坐标同时大于零的向量全体;

(iii) n 个坐标同时小于零的向量全体;

(iv) n 个坐标中前 r 个坐标大于零，后 $n-r$ 个小于零的向量全体.

2. 设 $\alpha_1, \cdots, \alpha_r$ 为 n 维线性空间 $\mathfrak{L}$ 中 r 个线性无关的向量. 记 $\mathfrak{L}_1$ 为由 $\alpha_1, \cdots, \alpha_r$ 线性生成的子空间，在 $\mathfrak{L}$ 中取 $\alpha_{r+1}, \cdots, \alpha_n$, 使得 $\alpha_1, \cdots, \alpha_r$, $\alpha_{r+1}, \cdots, \alpha_n$ 为 $\mathfrak{L}$ 的基. 则 $\mathfrak{L}$ 中向量属于 $\mathfrak{L}_1$ 当且仅当它关于基 $\alpha_1, \cdots, \alpha_n$ 之坐标为

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

3. 设 $\{\alpha_1, \cdots, \alpha_r\}$, $\{\beta_1, \cdots, \beta_s\}$ 为 n 维线性空间 $\mathfrak{L}$ 中两个向量组. 设后一组由前一组线性表出. 试证: 后一组的极大线性无关部分组的向量个数不多于前一组的极大线性无关部分组的向量个数.

4. 设 $\mathfrak{L}_1, \mathfrak{L}_2, \mathfrak{L}_3$ 为 n 维线性空间 $\mathfrak{L}$ 中三个子空间. 试证:

(i) $\mathfrak{L}_1 + \mathfrak{L}_1 + \cdots + \mathfrak{L}_1 = \mathfrak{L}_1$;

(ii) $\mathfrak{L}_1 \cap \mathfrak{L}_1 \cap \cdots \cap \mathfrak{L}_1 = \mathfrak{L}_1$;

(iii) $\mathfrak{L}_1 + \mathfrak{L} = \mathfrak{L} + \mathfrak{L}_1 = \mathfrak{L}$;

(iv) $\mathfrak{L}_1 \cap \mathfrak{L} = \mathfrak{L} \cap \mathfrak{L}_1 = \mathfrak{L}_1$;

(v) $(\mathfrak{L}_1 \cap \mathfrak{L}_2) + \mathfrak{L}_3 \subset (\mathfrak{L}_1 + \mathfrak{L}_3) \cap (\mathfrak{L}_2 + \mathfrak{L}_3)$. 何时相等?

(vi) $\mathfrak{L}_1 \cap \mathfrak{L}_2 + \mathfrak{L}_1 \cap \mathfrak{L}_3 \subset \mathfrak{L}_1 \cap (\mathfrak{L}_2 + \mathfrak{L}_3)$. 何时相等?

(vii) $\mathfrak{L}_1 \cap \mathfrak{L}_2 = \mathfrak{L}_1 + \mathfrak{L}_2$ 当且仅当 $\mathfrak{L}_1 = \mathfrak{L}_2$.

5. 设 $\mathfrak{L}_1$ 为 n 维线性空间 $\mathfrak{L}$ 的子空间，试证: 存在 $\mathfrak{L}$ 的子空间 $\mathfrak{L}_2$, 使得 $\mathfrak{L} = \mathfrak{L}_1 \oplus \mathfrak{L}_2$ 为子空间的直接和. 它唯一吗?

6. 设 $\mathfrak{L}_1,\mathfrak{L}_2,\cdots,\mathfrak{L}_t$ 为 n 维线性空间 $\mathfrak{L}$ 中 t 个子空间，试证：下面性质互相等价：

(1) $\dim(\mathfrak{L}_1+\mathfrak{L}_2+\cdots+\mathfrak{L}_t)=\sum\limits_{j=1}^{t}\dim\mathfrak{L}_j$，即 $\mathfrak{L}_1+\cdots+\mathfrak{L}_t$ 为子空间的直接和；

(2) $\mathfrak{L}_j\cap(\mathfrak{L}_1+\cdots+\mathfrak{L}_{j-1}+\mathfrak{L}_{j+1}+\cdots+\mathfrak{L}_t)=0,\ 1\leqslant j\leqslant t$;

(3) $\mathfrak{L}_1\cap\mathfrak{L}_2=0,(\mathfrak{L}_1+\mathfrak{L}_2)\cap\mathfrak{L}_3=0,\cdots,(\mathfrak{L}_1+\cdots+\mathfrak{L}_{t-1})\cap\mathfrak{L}_t=0$.

7. n 维线性空间 $\mathfrak{L}$ 中子空间 $\mathfrak{L}_0$ 称为**真子空间**，如果 $\mathfrak{L}_0\neq\mathfrak{L}$. 记 $\mathfrak{L}_1,\mathfrak{L}_2,\cdots,\mathfrak{L}_t$ 为 $\mathfrak{L}$ 中真子空间，试证：

(i) 存在 $\mathfrak{L}$ 中向量 α 使得 $\alpha\notin\mathfrak{L}_j,1\leqslant j\leqslant t$;

(ii) 在 $\mathfrak{L}$ 中存在基 $\alpha_1,\cdots,\alpha_n$，使得 $\alpha_j\notin\mathfrak{L}_j,j=1,2,\cdots,t$.

8. 设 $\mathfrak{L}_0,\mathfrak{L}_1,\cdots,\mathfrak{L}_t$ 为 n 维线性空间 $\mathfrak{L}$ 的子空间，且有 $\mathfrak{L}_0\subset\mathfrak{L}_1\cup\cdots\cup\mathfrak{L}_t$. 试证：存在 $1,2,\cdots,t$ 中指标 i，使得 $\mathfrak{L}_0\subset\mathfrak{L}_i$.

9. 试证：实对称方阵及 Hermite 方阵都有一个阶数等于秩的主子式不等于零.

§6.5 线性方程组求解的几何理论

定理 6.5.1 给定 m 个独立未知数 $x_1,x_2,\cdots,x_m$，n 个方程构成的线性方程组

$$Ax=\alpha.$$

记 $\alpha_1,\cdots,\alpha_m$ 为系数矩阵 A 的 m 个列向量，则线性方程组 $Ax=\alpha$ 有解的必要且充分条件为：α 是 $\alpha_1,\cdots,\alpha_m$ 的线性组合. 此即系数矩阵 A 与其增广矩阵 $\tilde{A}=(A,\alpha)$ 之秩相同.

证 充分性显然. 必要性证明如下：今 $Ax=\alpha$ 可改写为

$$\alpha=\sum_{j=1}^{m}x_j\alpha_j.$$

所以有解当且仅当 α 为 $\alpha_1,\cdots,\alpha_m$ 的线性组合. 此即 V_n 中向量组 $\alpha_1,\cdots,\alpha_m$ 的极大线性无关向量组也是 V_n 中向量组 $\alpha,\alpha_1,\cdots,\alpha_m$ 的极大线性无关向量组. 由定理 6.4.3，所以系数矩阵 A 和增广矩

阵 $\tilde{A}=(A,\alpha)$之秩相等. 证完.

定理 6.5.2 给定 m 个独立未知数 $x_1,x_2,\cdots,x_m$, n 个方程构成的线性方程组

$$Ax=0.$$

则它们的通解构成 V_m 中子空间, 维数等于 $m-\mathrm{rank}(A)$. 且基础解系为此子空间之基.

证 记 $n\times m$ 矩阵 A 的 m 个列向量为 $\alpha_1,\cdots,\alpha_m$. 则齐次线性方程组 $Ax=0$ 可改写为

$$x_1\alpha_1+\cdots+x_m\alpha_m=0.$$

在 V_n 中向量组 $\alpha_1,\cdots,\alpha_m$ 中任意取定极大线性无关部分组. 为讨论方便起见, 无妨设 α_1, $\cdots$, α_r 为极大线性无关部分组. 由定理 6.4.3, $r=\mathrm{rank}(A)$. 由极大线性无关部分组之定义可知

$$\alpha_j=\sum_{k=1}^{r}b_{jk}\alpha_k,\quad j=r+1,\cdots,m.$$

代回, 有

$$\begin{aligned}0=\sum_{j=1}^{m}x_j\alpha_j&=\sum_{j=1}^{r}x_j\alpha_j+\sum_{j=r+1}^{m}x_j\sum_{k=1}^{r}b_{jk}\alpha_k\\&=\sum_{j=1}^{r}\left(x_j+\sum_{k=r+1}^{m}x_kb_{kj}\right)\alpha_j.\end{aligned}$$

由于 $\alpha_1,\cdots,\alpha_r$ 线性无关, 所以有

$$x_j=-\sum_{k=r+1}^{m}x_kb_{kj},\quad 1\leqslant j\leqslant r.$$

记 $e_j\in V_m$, 其中 e_j 的第 j 个坐标为1, 其余坐标为零, $1\leqslant j\leqslant m$. 于是

$$\begin{aligned}x=\sum_{j=1}^{m}x_je_j&=\sum_{j=1}^{r}x_je_j+\sum_{j=r+1}^{m}x_je_j\\&=\sum_{j=1}^{r}\left(-\sum_{k=r+1}^{m}x_kb_{kj}\right)e_j+\sum_{k=r+1}^{m}x_ke_k\end{aligned}$$

$$=\sum_{k=r+1}^{m} x_k\left(e_k-\sum_{j=1}^{r} b_{kj}e_j\right).$$

记

$$\beta_k=e_k-\sum_{j=1}^{r} b_{kj}e_j,\quad r+1\leqslant k\leqslant m,$$

即有

$$x=\sum_{k=r+1}^{m} x_k\beta_k.$$

由于 $x_{r+1},\cdots,x_m$ 可任取. 这证明了通解构成由 $\beta_{r+1},\cdots,\beta_m$ 线性生成的子空间. 于是 $\beta_{r+1},\cdots,\beta_m$ 是基础解系. 余下证 $\beta_{r+1},\cdots,\beta_m$ 线性无关. 事实上,若存在常数 $b_{r+1},\cdots,b_m$, 使得 $\sum\limits_{j=r+1}^{m} b_j\beta_j=0$,此即

$$0=\sum_{k=r+1}^{m} b_k\beta_k=\sum_{k=r+1}^{m} b_k\left(e_k-\sum_{j=1}^{r} b_{kj},e_j\right)$$

$$=\sum_{j=1}^{r}\left(\sum_{k=r+1}^{m}(-b_kb_{kj})\right)e_j+\sum_{k=r+1}^{m} b_ke_k.$$

由于 $e_1,\cdots,e_m$ 为 V_m 之基,这证明了 $b_{r+1}=\cdots=b_m=0$. 所以 $\beta_{r+1},\cdots,\beta_m$ 线性无关. 证完.

习题 6.5

1. 记 $\mathfrak{L}_0=\{x_0\in V_m \mid ABx_0=0\}$, 其中 A,B 分别为 $p\times n$, $n\times m$ 矩阵. 试证: V_n 中子集

$$\mathfrak{L}_1=\{y\in V_n \mid y=Bx,\ \forall x\in\mathfrak{L}_0\}$$

为子空间,且有

$$\dim\mathfrak{L}_1=\operatorname{rank}(B)-\operatorname{rank}(AB).$$

用此试证明:对任意三个 n 阶方阵 A,B,C,则有

$$\operatorname{rank}(AB)+\operatorname{rank}(BC)\leqslant\operatorname{rank}(B)+\operatorname{rank}(ABC).$$

试另外用矩阵在相抵下的标准形来证明这个不等式，且比较何者简单.

2. 符号同第1题. 试证：$\operatorname{rank}(AB)=\operatorname{rank}(B)$ 当且仅当由 $ABx=0$ 可推出 $Bx=0$.

3. 设 A,B,C 分别为 $m\times n, n\times p, p\times q$ 矩阵. 试用本节理论证明：若 $\operatorname{rank}(B)=\operatorname{rank}(AB)$，则有 $\operatorname{rank}(BC)=\operatorname{rank}(ABC)$.

4. 设 A 和 B 分别为 $s\times n$ 及 $t\times n$ 矩阵. 记 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 分别为 n 维线性空间 V_n 中子空间，且 $\mathfrak{L}_1$ 为齐次线性方程组 $Ax=0$ 之解空间，$\mathfrak{L}_2$ 为齐次线性方程组 $Bx=0$ 之解空间. 试构造一个齐次线性方程组，使得它的解空间为 V_n 中子空间 $\mathfrak{L}_1\cap\mathfrak{L}_2$；另外构造一个齐次线性方程组，使得它的解空间为 $\mathfrak{L}_1+\mathfrak{L}_2$.

5 设 A,B,C 分别为 $m\times n, n\times p, n\times q$ 矩阵，且

$$\operatorname{rank}(A)+\operatorname{rank}(B)=n,\quad AB=0,\quad AC=0.$$

试用本节理论来证明：存在 $p\times q$ 矩阵 D，使得 $C=BD$，且唯一存在 $p\times q$ 矩阵 D 的必要且充分条件为 $\operatorname{rank}(B)=p$.

第七章 线 性 变 换

§7.1 线 性 变 换

定义 线性空间 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 内的映射称为**零映射**，如果它将 $\mathfrak{L}_1$ 中每个向量都映为 $\mathfrak{L}_2$ 中的零向量. 零映射用通常的数 0 符号来表示，这样并不会产生混淆的. 线性空间 $\mathfrak{L}$ 到自身内的零映射称为零变换.

定义 线性空间 $\mathfrak{L}$ 到自身上的映射称为**恒等变换**，如果它将 $\mathfrak{L}$ 中任一向量都映为自己. 恒等变换总用符号 id 来表示，确切地可表为 $id_{\mathfrak{L}}$.

定义 线性空间 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 上的映射 σ 称为**可逆的**，如果存在 $\mathfrak{L}_2$ 到 $\mathfrak{L}_1$ 上的映射 τ，使得

$$\tau\sigma = id_{\mathfrak{L}_1}$$

这时 τ 称为 σ 的**逆映射**，记作 σ^{-1}.

引理 7.1.1 σ 为线性空间 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 上的可逆映射当且仅当 σ 为 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 上的一一映射. 这时逆映射 τ 唯一存在，且有

$$\sigma\tau = id_{\mathfrak{L}_2}.$$

证 若 σ 为 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 上的可逆映射，即存在 $\mathfrak{L}_2$ 到 $\mathfrak{L}_1$ 上的映射 τ，使得 $\tau\sigma = id_{\mathfrak{L}_1}$，下面来证 σ 为一一映射. 事实上，任取 $\alpha, \beta \in \mathfrak{L}_1$，使得 $\sigma(\alpha) = \sigma(\beta) = \delta \in \mathfrak{L}_2$，于是 $\tau(\delta) = \tau\sigma(\alpha) = \tau\sigma(\beta)$. 由于 $\tau\sigma = id_{\mathfrak{L}_1}$，因此 $\alpha = \beta$，这证明了 σ 为一一映射. 反之，若 σ 为 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 上的一一映射，$\sigma(\mathfrak{L}_1) = \mathfrak{L}_2$. 任取 $\beta \in \mathfrak{L}_2$，则存在 $\alpha \in \mathfrak{L}_1$，使得 $\sigma(\alpha) = \beta$. 由 σ 一一，可知 α 唯一存在，因此 $\tau: \beta \to \alpha$ 为 $\mathfrak{L}_2$ 到 $\mathfrak{L}_1$ 内

的映射，且 $\alpha=\tau(\beta)=\tau(\sigma(\alpha))=\tau\sigma(\alpha)$. 但是 $\sigma(\mathfrak{L}_1)=\mathfrak{L}_2$ 告诉我们，当 β 遍历 $\mathfrak{L}_2$ 时，α 遍历 $\mathfrak{L}_1$，所以 τ 为到上的映射，即 $\alpha=\tau\sigma(\alpha)$，$\forall\alpha\in\mathfrak{L}_1$. 这证明了 $\tau\sigma=id_{\mathfrak{L}_1}$，即 σ 可逆，且 τ 为 σ 的逆映射. 下面证逆映射唯一. 今若另有 $\mathfrak{L}_2$ 到 $\mathfrak{L}_1$ 上的映射 τ_1 适合 $\tau_1\sigma=id_{\mathfrak{L}_1}$，于是 $\tau\sigma(\alpha)=\tau_1\sigma(\alpha)=\alpha$，$\forall\alpha\in\mathfrak{L}_1$. 由 $\sigma(\mathfrak{L}_1)=\mathfrak{L}_2$，所以任取 $\beta\in\mathfrak{L}_2$，则有 $\tau(\beta)=\tau_1(\beta)$，即证明了 $\tau=\tau_1$. 最后证 $\sigma\tau=id_{\mathfrak{L}_2}$. 今已知 $\tau\sigma=id_{\mathfrak{L}_1}$，于是任取 $\alpha\in\mathfrak{L}_1$，有 $\tau\sigma(\alpha)=\alpha$，因此 $(\sigma\tau)\sigma(\alpha)=\sigma(\tau\sigma(\alpha))=\sigma(\alpha)$. 由 $\sigma(\mathfrak{L}_1)=\mathfrak{L}_2$ 可知 $\sigma\tau(\beta)=\beta$，$\forall\beta\in\mathfrak{L}_2$，这证明了 $\sigma\tau=id_{\mathfrak{L}_2}$. 证完.

显然，零映射和恒等变换都是线性的，但是可逆映射一般不是线性映射.

设 σ 为线性空间 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 内的线性映射. 由定义可知下列性质成立：

(1) 任取数 $b_1,b_2,\cdots,b_r$，任取 $\beta_1,\beta_2,\cdots,\beta_r\in\mathfrak{L}_1$，则有

$$\sigma\left(\sum_{j=1}^{r}b_j\beta_j\right)=\sum_{j=1}^{r}b_j\sigma(\beta_j).$$

(2) 若 $\mathfrak{L}_1$ 中向量 $\beta_1,\cdots,\beta_r$ 线性相关，则 $\mathfrak{L}_2$ 中向量 $\sigma(\beta_1),\cdots,\sigma(\beta_r)$ 也线性相关，因此若 $\sigma(\beta_1),\cdots,\sigma(\beta_r)$ 线性无关，则 $\beta_1,\cdots,\beta_r$ 也线性无关.

注意，$\beta_1,\cdots\beta_r$ 线性无关，推不出 $\sigma(\beta_1),\cdots,\sigma(\beta_r)$ 线性无关，又 $\sigma(\beta_1),\cdots,\sigma(\beta_r)$ 线性相关，推不出 $\beta_1,\cdots,\beta_r$ 线性相关.

(3) $\sigma(\mathfrak{L}_1)=\{\sigma(\alpha)\mid\forall\alpha\in\mathfrak{L}_1\}$ 为 $\mathfrak{L}_2$ 的子空间，称为 σ 的**象空间**.

(4) $\ker(\sigma)=\sigma^{-1}(0)=\{\beta\in\mathfrak{L}_1\mid\sigma(\beta)=0\}$ 为 $\mathfrak{L}_1$ 的子空间，称为 σ 的**核**.

利用 σ 的核，我们可以描述**完全原象**.

(5) 任取 $\beta\in\sigma(\mathfrak{L}_1)\subset\mathfrak{L}_2$，则 β 的完全原象

$$\sigma^{-1}(\beta)=\{\alpha\in\mathfrak{L}_1 \mid \sigma(\alpha)=\beta\}$$

非空，在其中任意取定向量 α_0，则

$$\sigma^{-1}(\beta)=\alpha_0+\ker(\sigma)=\{\alpha_0+\alpha\in\mathfrak{L}_1 \mid \forall\alpha\in\ker(\sigma)\}.$$

定理 7.1.1 在线性空间 $\mathfrak{L}_1$ 及 $\mathfrak{L}_2$ 中分别取定基 $\{\alpha_1,\cdots,\alpha_m\}$，$\{\beta_1,\cdots,\beta_n\}$，则 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 内的线性映射 σ 有

$$\sigma(\alpha_i)=\sum_{j=1}^{n}a_{ji}\beta_j,\quad i=1,2,\cdots,m.$$

于是 σ 在基 $\{\alpha_1,\cdots,\alpha_m\}$ 和 $\{\beta_1,\cdots,\beta_n\}$ 下和 $n\times m$ 矩阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

间有一个一一对应关系，所以 $n\times m$ 矩阵 A 称为线性映射 σ 在基 $\{\alpha_1,\cdots,\alpha_m\}$，$\{\beta_1,\cdots,\beta_n\}$ 下的**矩阵表示**

证 由 $\beta_1,\cdots,\beta_n$ 为基，所以 $\sigma\to A$ 为映射，它一一. 事实上，若 $\tau\to A$，τ 为 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 内的线性映射，则 $\sigma(\alpha_i)=\sum_{j=1}^{n}a_{ji}\beta_j$，$\tau(\alpha_i)=\sum_{j=1}^{n}a_{ji}\beta_j$，$1\leqslant i\leqslant m$. 这证明了

$$\sigma(\alpha_i)=\tau(\alpha_i),\quad i=1,2,\cdots,m.$$

对 $\mathfrak{L}_1$ 中任一元 $\alpha=\sum_{i=1}^{m}a_i\alpha_i$，则 $\sigma(\alpha)=\sum_{i=1}^{m}a_i\sigma(\alpha_i)=\sum_{i=1}^{m}a_i\tau(\alpha_i)=\tau(\alpha)$. 这证明了 $\sigma=\tau$，所以映射 $\sigma\to A$ 为一一映射，最后证这是线性映射集到 $n\times m$ 矩阵集上的映射. 事实上，任取 $n\times m$ 矩阵 $B=(b_{ij})$，作对应

$$\alpha_i\to\sum_{j=1}^{n}b_{ji}\beta_j,\quad i=1,2,\cdots,m$$

以及任取 $\alpha\in\mathfrak{L}_1$，作对应

$$\alpha=\sum_{i=1}^{m}a_i\alpha_i\rightarrow\sum_{i=1}^{m}\sum_{j=1}^{n}a_ib_{ji}\beta_j$$

它给出了 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 内的单值映射. 由直接验证可知它是线性映射，记作 ξ，这证明了：任给 $n\times m$ 矩阵 B，存在线性映射 $\xi:\mathfrak{L}_1\rightarrow\mathfrak{L}_2$，使得 ξ 有矩阵表示 B，所以证明了定理. 证完.

由于线性映射的矩阵表示和基底选取有关，很自然地，当选取不同基时，矩阵表示如何改变，便是一个重要问题.

在 $\mathfrak{L}_1$ 中取定基 $\{\alpha_1, \cdots, \alpha_m\}$，$\{\alpha'_1, \cdots, \alpha'_m\}$，在 $\mathfrak{L}_2$ 中取定基 $\{\beta_1,\cdots,\beta_n\}$，$\{\beta'_1,\cdots,\beta'_n\}$. 于是有基变换公式

$$\alpha'_i=\sum_{j=1}^{m}p_{ji}\alpha_j,\quad 1\leqslant i\leqslant m,$$

$$\beta'_i=\sum_{j=1}^{n}q_{ji}\beta_j,\quad 1\leqslant i\leqslant n.$$

它们分别对应非异方阵

$$P=\begin{pmatrix}p_{11}&\cdots&p_{1m}\\\vdots&&\vdots\\p_{m1}&\cdots&p_{mm}\end{pmatrix},\qquad Q=\begin{pmatrix}q_{11}&\cdots&q_{1n}\\\vdots&&\vdots\\q_{n1}&\cdots&q_{nn}\end{pmatrix}.$$

现在给定 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 内的线性映射 σ，于是有

$$\sigma(\alpha_i)=\sum_{j=1}^{n}a_{ji}\beta_j,\quad 1\leqslant i\leqslant m,$$

$$\sigma(\alpha'_i)=\sum_{j=1}^{n}b_{ji}\beta_j',\quad 1\leqslant i\leqslant m.$$

其中

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1m}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nm}\end{pmatrix},\qquad B=\begin{pmatrix}b_{11}&\cdots&b_{1m}\\\vdots&&\vdots\\b_{n1}&\cdots&b_{nm}\end{pmatrix}$$

分别为线性映射 σ 在基 $\{\alpha_1,\cdots,\alpha_n\}$，$\{\beta_1,\cdots,\beta_n\}$ 以及 $\{\alpha'_1,\cdots,\alpha'_m\}$，

$\{\beta'_1,\cdots,\beta'_n\}$下的矩阵表示.

定理 7.1.2 符号同上,则有

$$B=Q^{-1}AP.$$

证 今

$$\sigma(\alpha'_i)=\sigma\left(\sum_{j=1}^{m}p_{ji}\alpha_j\right)=\sum_{j=1}^{m}p_{ji}\sigma(\alpha_j)=\sum_{j=1}^{m}p_{ji}\sum_{k=1}^{n}a_{kj}\beta_k$$
$$=\sum_{k=1}^{n}\left(\sum_{j=1}^{m}a_{kj}p_{ji}\right)\beta_k,$$

另一方面,

$$\sigma(\alpha'_i)=\sum_{j=1}^{n}b_{ji}\beta'_j=\sum_{j=1}^{n}b_{ji}\sum_{k=1}^{n}q_{kj}\beta_k=\sum_{k=1}^{n}\left(\sum_{j=1}^{n}q_{kj}b_{ji}\right)\beta_k.$$

这证明了

$$\sum_{j=1}^{m}a_{kj}p_{ji}=\sum_{j=1}^{n}q_{kj}b_{ji},\quad 1\leqslant i\leqslant m,1\leqslant k\leqslant n.$$

即有 $AP=QB$. 由 Q 非异,所以有 $B=Q^{-1}AP$. 证完.

由相抵下标准形理论,立即有

定理7.1.3 设 σ 为线性空间 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 内的线性映射,则在 $\mathfrak{L}_1$ 及 $\mathfrak{L}_2$ 中分别存在基 $\{\alpha_1,\cdots,\alpha_m\}$, $\{\beta_1,\cdots,\beta_n\}$, 使得 σ 关于这两组基的矩阵表示为标准形

$$\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix},$$

因此有

$$\sigma(\alpha_i)=\beta_i,\quad 1\leqslant i\leqslant r,\quad \sigma(\alpha_i)=0,\quad r+1\leqslant i\leqslant m.$$

于是 σ 之象空间 $\sigma(\mathfrak{L}_1)$ 为 $\mathfrak{L}_2$ 中以 $\beta_1,\cdots,\beta_r$ 为基的子空间, 又 σ 之核空间 $\ker(\sigma)$ 是 $\mathfrak{L}_1$ 中以 $\alpha_{r+1},\cdots,\alpha_m$ 为基的子空间.

定理 7.1.4 设 σ 为线性空间 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 内的线性映射,则有

$$\dim\mathfrak{L}_1=\dim(\sigma(\mathfrak{L}_1))+\dim(\ker(\sigma)),$$

且$\dim\sigma(\mathfrak{L}_1)$为$\sigma$之任一矩阵表示之秩，它也称为**线性映射$\sigma$之秩**. 所以$\sigma$为到上的线性同构当且仅当$\sigma(\mathfrak{L}_1)=\mathfrak{L}_2$，且$\ker(\sigma)=0$.

证 由定理7.1.3，所以$\dim(\sigma(\mathfrak{L}_1))=r$，$\dim(\ker(\sigma))=m-r$. 于是$\dim(\sigma(\mathfrak{L}_1))+\dim(\ker(\sigma))=m=\dim\mathfrak{L}_1$. 又$\dim(\sigma(\mathfrak{L}_1))=r=\operatorname{rank}(A)$，其中$A$为$\sigma$的任一矩阵表示. 证完.

下面考虑线性映射σ的坐标表达式，设σ为线性空间$\mathfrak{L}_1$到$\mathfrak{L}_2$内的线性映射. 在$\mathfrak{L}_1$及$\mathfrak{L}_2$中分别取基$\{\alpha_1,\cdots,\alpha_m\}$，$\{\beta_1,\cdots,\beta_n\}$. 记$A=(a_{ij})$为$\sigma$在上述基下的矩阵表示，即有

$$\sigma(\alpha_i)=\sum_{j=1}^{n}a_{ji}\beta_j,\quad i=1,2,\cdots,m.$$

现在在$\mathfrak{L}_1$中任取一向量$\alpha=\sum_{i=1}^{m}x_i\alpha_i$，则$\sigma(\alpha)$为$\mathfrak{L}_2$中向量，即有$\sigma(\alpha)=\sum_{j=1}^{n}y_j\beta_j$. 所以

$$\sum_{j=1}^{n}y_j\beta_j=\sigma(\alpha)=\sum_{i=1}^{m}x_i\sigma(\alpha_i)=\sum_{i=1}^{m}\sum_{j=1}^{n}x_ia_{ji}\beta_j,$$

这证明了

$$y_j=\sum_{i=1}^{m}a_{ji}x_i,\quad j=1,2,\cdots,n.$$

记

$$x=\begin{pmatrix}x_1\\ \vdots\\ x_m\end{pmatrix}\in V_m,\qquad y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\in V_n,$$

则上式可写成矩阵形式

$$y=Ax.$$

下面考虑线性变换，即考虑线性空间$\mathfrak{L}$到自身内的线性映射，这时我们可以用下面办法来定义线性变换$\mathscr{A}$的方阵表示. 我们用

定理 7.1.1 的结果，立即有

定理 7.1.5 在 n 维线性空间 $\mathfrak{L}$ 中取基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 则 $\mathfrak{L}$ 上线性变换 σ 有

$$\sigma(\alpha_i)=\sum_{j=1}^{n} a_{ji}\alpha_j, \qquad 1\leqslant i\leqslant n.$$

于是线性变换 σ 在基 $\{\alpha_1,\cdots,\alpha_n\}$ 下有方阵表示

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

且对应 $\sigma\to A$ 为 $\mathfrak{L}$ 上所有线性变换到所有 n 阶方阵集合上的一一对应，这时对 $\mathfrak{L}$ 中任一向量 $\alpha=\sum_{i=1}^{n} x_i\alpha_i$，记 $\sigma(\alpha)=\sum_{j=1}^{n} y_j\alpha_j$，且记

$$x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad y=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\in V_n,$$

则有

$$y=Ax.$$

所以它称为线性变换 σ 的**坐标表达式**

用定理 7.1.2 之结果，立即有

定理 7.1.6 在 n 维线性空间 $\mathfrak{L}$ 中取定两组基 $\{\alpha_1,\cdots,\alpha_n\}$ 和 $\{\beta_1,\cdots,\beta_n\}$. 记基变换

$$\beta_i=\sum_{j=1}^{n} p_{ji}\alpha_j, \qquad 1\leqslant i\leqslant n$$

对应的 n 阶非异方阵为

$$P=\begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}.$$

则 $\mathfrak{L}$ 上同一线性变换 $\mathscr{A}$ 在不同基下的方阵表示 A 及 B 有关系

$$B=P^{-1}AP.$$

定义 n 阶方阵 A 和 B 称为**相似的**，如果存在 n 阶非异方阵 P，使得

$$B=P^{-1}AP.$$

由这个定义，我们有

定理 7.1.7 n 维线性空间 $\mathfrak{L}$ 上线性变换在不同基下的方阵表示互相相似.

引理 7.1.2 n 阶方阵间的相似关系为等价关系.

证 显然 A 和 A 相似，这证明了反身性. 今若 A 和 B 相似，即存在非异方阵 P，使得 $B=P^{-1}AP$，于是 $A=PBP^{-1}$. 这证明了 B 和 A 相似，即证明了对称性. 最后，若 A 和 B 相似，B 和 C 相似，即存在 n 阶非异方阵 P 及 Q 使得 $B=P^{-1}AP$，$C=Q^{-1}BQ$，于是有 $C=Q^{-1}(P^{-1}AP)Q=(PQ)^{-1}A(PQ)$. 这证明了 A 和 C 相似，即证明了传递性. 证完.

定理 7.1.8 设 $\mathscr{A}$ 为 n 维线性空间 $\mathfrak{L}$ 上线性变换，则下面性质互相等价：

(1) $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{L}$ 上的线性同构，即 $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{L}$ 上的一一线性变换；

(2) $\mathscr{A}(\mathfrak{L})=\mathfrak{L}$，即 $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{L}$ 上的线性变换；

(3) $\ker(\mathscr{A})=0$；

(4) 在 $\mathfrak{L}$ 中取定基 $\alpha_1,\cdots,\alpha_n$，记 A 为 $\mathscr{A}$ 在基 $\alpha_1,\cdots,\alpha_n$ 下的方阵表示，则 A 可逆.

证 先由(1)推(2)，事实上，由于线性同构定义要求 $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{L}$ 上的一一线性变换，所以有 $\mathscr{A}(\mathfrak{L})=\mathfrak{L}$. 再由(2)推(3). 这是因为由定理 7.1.4 有 $\dim\mathfrak{L}=\dim(\mathscr{A}(\mathfrak{L}))+\dim(\ker(\mathscr{A}))$、再由 $\mathscr{A}(\mathfrak{L})=\mathfrak{L}$，便证明了 $\dim(\ker(\mathscr{A}))=0$，所以 $\ker(\mathscr{A})=0$. 再由

(3)推(4)，事实上，由 $\ker(\mathscr{A})=0$，所以 $\dim\Omega=\dim(\mathscr{A}(\Omega))+\dim(\ker(\mathscr{A}))=\dim\mathscr{A}(\Omega)$．而 $\dim\mathscr{A}(\Omega)$ 等于 $\mathscr{A}$ 的方阵表示的秩，所以等于 $\dim\Omega=n$，即$\mathscr{A}$在基 $\alpha_1,\cdots,\alpha_n$ 下的方阵表示 A 非异．

最后证由(4)可推出(1)．今在基 $\alpha_1,\cdots,\alpha_n$ 下线性变换$\mathscr{A}$的方阵表示 A 为可逆方阵．记 $A=(a_{ij})$，$A^{-1}=(b_{ij})$，于是可以定义 Ω 上线性变换 $\mathscr{L}$，使得它在基 $\alpha_1,\cdots,\alpha_n$下的方阵表示为 A^{-1}．令

$$\mathscr{A}(\alpha_i)=\sum_{j=1}^{n}a_{ji}\alpha_j,\qquad \mathscr{L}(\alpha_i)=\sum_{j=1}^{n}b_{ji}\alpha_j,\qquad 1\leqslant i\leqslant n.$$

于是$\sum_{i=1}^{n}a_{ik}\mathscr{L}(\alpha_i)=\sum_{i=1}^{n}a_{ik}\sum_{j=1}^{n}b_{ji}\alpha_j=\sum_{j=1}^{n}\delta_{jk}\alpha_j=\alpha_k$，$1\leqslant k\leqslant n$，此即 $\mathscr{L}(\Omega)=\Omega$．又

$$\mathscr{A}(\mathscr{L}(\alpha_i))=\mathscr{A}(\sum_{j=1}^{n}b_{ji}\alpha_j)=\sum_{j=1}^{n}b_{ji}\mathscr{A}(\alpha_j)=\sum_{j=1}^{n}b_{ji}\sum_{k=1}^{n}a_{kj}\alpha_k$$
$$=\sum_{k=1}^{n}\delta_{ki}\alpha_k=\alpha_i,$$

$$(\mathscr{L}\mathscr{A}(\alpha_i))=\mathscr{L}(\sum_{j=1}^{n}a_{ji}\alpha_j)=\sum_{j=1}^{n}a_j\mathscr{L}_i(\alpha_j)=\sum_{j=1}^{n}a_{ji}\sum_{k=1}^{n}b_{kj}\alpha_k$$
$$=\sum_{k=1}^{n}\delta_{ki}\alpha_k=\alpha_i,\qquad i=1,2,\cdots,n.$$

这证明了$\mathscr{A}(\mathscr{L}(\alpha))=\mathscr{L}(\mathscr{A}(\alpha))=\alpha$，$\forall\alpha\in\Omega$．所以$\mathscr{A}\mathscr{L}=id$，$\mathscr{L}\mathscr{A}=id$，此即 $\mathscr{L}=\mathscr{A}^{-1}$．所以$\mathscr{A}$为 Ω 到 Ω 上的一一线性变换，即为线性同构．证完．

定义 在 n 维线性空间 Ω 上任给线性变换$\mathscr{A}$及$\mathscr{L}$，以及纯量 a，则

(1) 映射

$$\alpha\to\mathscr{A}(\alpha)+\mathscr{L}(\alpha),\quad \forall\alpha\in\Omega$$

仍为 Ω 上线性变换，称为线性变换$\mathscr{A}$和$\mathscr{L}$的和，记作 $\mathscr{A}+\mathscr{L}$．于

是有

$$(\mathscr{A}+\mathscr{L})(\alpha)=\mathscr{A}(\alpha)+\mathscr{L}(\alpha),\quad \forall\alpha\in\Omega.$$ ▎

(2) 映射

$$\alpha\to\mathscr{A}(\mathscr{L}(\alpha)),\quad \forall\alpha\in\Omega$$

仍为 Ω 上线性变换，称为线性变换$\mathscr{A}$和$\mathscr{L}$的乘积，记作$\mathscr{A}\mathscr{L}$. 于是有

$$(\mathscr{A}\mathscr{L})(\alpha)=\mathscr{A}(\mathscr{L}(\alpha)),\quad \forall\alpha\in\Omega.$$

(3) 映射

$$\alpha\to a\mathscr{A}(\alpha),\quad \forall\alpha\in\Omega$$

仍为 Ω 上线性变换，称为线性变换$\mathscr{A}$和纯量 a 的**纯量乘积**，记作 $a\mathscr{A}$. 于是有

$$(a\mathscr{A})(\alpha)=a(\mathscr{A}(\alpha)),\quad \forall\alpha\in\Omega.$$

定理 7.1.9 在 n 维线性空间 Ω 中取定一组基 $\{\alpha_1,\cdots,\alpha_n\}$. 记 n 阶方阵 A 和 B 分别为 Ω 上线性变换$\mathscr{A}$和$\mathscr{L}$在基$\{\alpha_1,\cdots,\alpha_n\}$下的方阵表示, a 为数, 则线性变换 $\mathscr{A}+\mathscr{L}$, $\mathscr{A}\mathscr{L}$, $a\mathscr{A}$在基$\{\alpha_1,\cdots,\alpha_n\}$下的方阵表示分别为 $A+B$, AB, aA.

证 记 $A=(a_{ij})$, $B=(b_{ij})$, 于是

$$\mathscr{A}(\alpha_i)=\sum_{j=1}^{n}a_{ji}\alpha_j,\qquad \mathscr{L}(\alpha_i)=\sum_{j=1}^{n}b_{ji}\alpha_j,\qquad 1\leqslant i\leqslant n.$$

因此

$$(\mathscr{A}+\mathscr{L})(\alpha_i)=\mathscr{A}(\alpha_i)+\mathscr{L}(\alpha_i)=\sum_{j=1}^{n}(a_{ji}+b_{ji})\alpha_j,$$

$$(\mathscr{A}\mathscr{L})(\alpha_i)=\mathscr{A}(\mathscr{L}(\alpha_i))=\mathscr{A}\Big(\sum_{j=1}^{n}b_{ji}\alpha_j\Big)=\sum_{j=1}^{n}\sum_{k=1}^{n}b_{ji}a_{kj}\alpha_k,$$

$$(a\mathscr{A})(\alpha_i)=a(\mathscr{A}(\alpha_i))=a\sum_{j=1}^{n}a_{ji}\alpha_j=\sum_{j=1}^{n}(aa_{ji})\alpha_j,$$

$i=1,2,\cdots,n$. 于是 $\mathscr{A}+\mathscr{L}$, $\mathscr{A}\mathscr{L}$, $a\mathscr{A}$的方阵表示分别为 $A+B$,

AB, aA. 证完.

上面的定理告诉我们，方阵的乘法，加法和纯量乘积的定义都是极其自然的，有明确的几何意义.

习题 7.1

1. 试证：线性变换将子空间变为子空间，非异线性变换将 r 维子空间变为 r 维子空间.

2. 设 $\mathscr{A}$ 为 n 维线性空间 Ω 上连续变换，即在 Ω 中任意取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$，任取 Ω 中向量 $\alpha=\sum_{j=1}^{n} x_j\alpha_j$，则

$$\mathscr{A}(\alpha)=\sum_{j=1}^{n} f_j(x_1, \cdots, x_n)\alpha_j,$$

其中 $f_1, \cdots, f_n$ 为连续函数. 设 $\mathscr{A}$ 还有性质：

$$\mathscr{A}(\mathscr{A}_1+\mathscr{A}_2)=\mathscr{A}\mathscr{A}_1+\mathscr{A}\mathscr{A}_2$$

对 Ω 上任两线性变换 $\mathscr{A}_1, \mathscr{A}_2$ 成立，则 $\mathscr{A}$ 也是线性变换.

3. 在 n 阶方阵全体构成的线性空间 M_n 上引进映射 $\mathscr{A}$：

$$X \longrightarrow AXB+CX+XD,$$

其中 A, B, C, D 都是给定的 n 阶方阵. 试证：$\mathscr{A}$ 为 M_n 上线性变换. 设 $C=D=0$，则 $\mathscr{A}$ 非异当且仅当 $\det A \det B \neq 0$. 设 $A=0, D=-C$，试给出这个线性变换的方阵表示.

4. 在次数不超过 n 的多项式全体构成的 $n+1$ 维线性空间 Ω 中引进映射

$$f(x) \longrightarrow f(x+1)-f(x), \quad \forall f \in P_n[X]$$

试回答下列问题：这个映射是不是线性变换.

5. 记 Ω 为所有多项式 $f(x, y)$ 构成之集合，其中 f 为关于独立未知数 x 和 y 的多项式，且次数都不超过 n. 试在 Ω 中给出一组基，并写出 Ω 的维数. 又映射

$$\sigma: f(x, y) \longrightarrow (x^2+y^2) f(x, y)$$

有性质：

$$\tau=(\partial_x\partial_x+\partial_y\partial_y)\cdot\sigma-\sigma\cdot(\partial_x\partial_x+\partial_y\partial_y)$$

为 $\mathfrak{L}$ 上线性变换. 试求线性变换 τ 在基$\frac{1}{j!k!}x^jy^k$, $0\leqslant j$, $k\leqslant n$ 下的方阵表示.

§7.2 不变子空间和Jordan标准形

定义 在 n 维线性空间 $\mathfrak{L}$ 上给定线性变换 $\mathscr{A}$. $\mathfrak{L}$ 中子空间 $\mathfrak{L}_1$ 称为线性变换$\mathscr{A}$的**不变子空间**,如果

$$\mathscr{A}(\mathfrak{L}_1)\subset\mathfrak{L}_1.$$

显然,给定线性变换$\mathscr{A}$,则 $\mathfrak{L}$ 的子空间

$$\mathfrak{L},\ 0,\ \mathscr{A}(\mathfrak{L}),\ \ker(\mathscr{A})$$

都是$\mathscr{A}$的不变子空间. 且我们有

引理 7.2.1 设$\mathscr{A}$为 n 维线性空间 $\mathfrak{L}$ 上线性变换, $\mathfrak{L}_1$ 为 $\mathfrak{L}$ 中关于线性变换$\mathscr{A}$的不变子空间. 在 $\mathfrak{L}_1$ 中取定基$\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$,于是 $\mathfrak{L}$ 中可取基$\{\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n\}$. 且在 $\mathfrak{L}$ 中这组基下, 线性变换 $\mathscr{A}$ 的方阵表示为

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1r} & a_{1,r+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} & a_{r,r+1} & \cdots & a_{rn} \\ 0 & \cdots & 0 & a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{n,r+1} & \cdots & a_{nn} \end{pmatrix}=\begin{pmatrix} A_1^{(r)} & A_3^{(r,n-r)} \\ 0 & A_2^{(n-r)} \end{pmatrix},$$

且由 $\mathscr{A}(\mathfrak{L}_1)\subset\mathfrak{L}_1$,所以$\mathscr{A}$限制在 $\mathfrak{L}_1$ 上,为 $\mathfrak{L}_1$ 上线性变换,记作

$$\mathscr{A}_1=\mathscr{A}|_{\mathfrak{L}_1}.$$

这时在 $\mathfrak{L}_1$ 的基$\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$下线性变换 $\mathscr{A}_1$ 之方阵表示为 A_1.

证 今在基 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$ 下,线性变换$\mathscr{A}$之方阵表示 $A=(a_{ij})$有 $\mathscr{A}(\alpha_i)=\sum\limits_{j=1}^{n}a_{ji}\alpha_j$, $1\leqslant i\leqslant n$. 但是$\alpha_1,\cdots,\alpha_r$ 为 $\mathfrak{L}_1$ 之基, 由$\mathscr{A}(\mathfrak{L}_1)\subset\mathfrak{L}_1$, 所以 $\mathscr{A}(\alpha_1),\cdots,\mathscr{A}(\alpha_r)$为 $\alpha_1,\cdots,\alpha_r$ 的线性组

合. 这证明了 $a_{ji}=0, i=1,2,\cdots,r, j=r+1,r+2,\cdots,n$. 且

$$\mathscr{A}(\alpha_i)=\sum_{j=1}^{r}a_{ji}\alpha_j, \qquad i=1,2,\cdots,r.$$

所以$\mathscr{A}|_{\mathfrak{L}_1}$在 $\alpha_1,\cdots,\alpha_r$ 下之方阵表示为 A_1. 证完.

引理 7.2.2 设$\mathscr{A}$为 n 维线性空间 $\mathfrak{L}$ 上线性变换. 设在 $\mathfrak{L}$ 中存在一组基 $\alpha_1,\cdots,\alpha_n$, 使得$\mathscr{A}$在这组基下有方阵表示

$$A=\begin{pmatrix} A_1^{(r)} & A_3^{(r,n-r)} \\ 0 & A_2^{(n-r)} \end{pmatrix}.$$

则以 $\alpha_1,\cdots,\alpha_r$ 为基之子空间 $\mathfrak{L}_1$ 为线性变换$\mathscr{A}$的不变子空间, 且$\mathscr{A}|_{\mathfrak{L}_1}$在基$\{\alpha_1,\cdots,\alpha_r\}$下的方阵表示为 A_1.

证明显然.

由于 n 维线性空间 $\mathfrak{L}$ 上线性变换在不同基下的方阵表示互相相似, 所以对 n 阶方阵 A, 找 n 阶非异方阵 P 使得 A 相似于准上三角方阵 $P^{-1}AP$ 的问题, 变为首先在 $\mathfrak{L}$ 中寻找线性变换$\mathscr{A}$的不变子空间. 我们有

引理 7.2.3 设$\mathscr{A}$为 n 维线性空间 $\mathfrak{L}$ 上线性变换, 设 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 为$\mathscr{A}$的不变子空间, 且有直接和 $\mathfrak{L}=\mathfrak{L}_1\oplus\mathfrak{L}_2$. 在 $\mathfrak{L}_1$ 中任取一组基 $\alpha_1,\cdots,\alpha_r$, 在$\mathfrak{L}_2$ 中任取一组基 $\alpha_{r+1},\cdots,\alpha_n$, 则 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$ 为 $\mathfrak{L}$ 的基. 且在这组基下, 线性变换$\mathscr{A}$的方阵表示 A 为准对角形

$$A=\mathrm{diag}(A_1^{(r)}, A_2^{(n-r)}),$$

其中 A_i 为 $\mathscr{A}|_{\mathfrak{L}_i}$ 在 $\mathfrak{L}_i$ 的上述基下的方阵表示, $i=1,2$.

证 由$\mathscr{A}(\mathfrak{L}_i)\subset\mathfrak{L}_i, i=1,2$, 于是

$$\mathscr{A}(\alpha_i)=\sum_{j=1}^{r}a_{ji}\alpha_j, \qquad i=1,2,\cdots,r,$$

$$\mathscr{A}(\alpha_i)=\sum_{j=r+1}^{n}a_{ji}\alpha_j, \qquad i=r+1,\cdots,n.$$

所以记$\mathscr{A}|_{\mathfrak{L}_1}$在基$\alpha_1,\cdots,\alpha_r$下的方阵表示为$A_1$, $\mathscr{A}|_{\mathfrak{L}_2}$在基 α_{r+1},

$\cdots,\alpha_n$ 下的方阵表示为A_2，则$\mathscr{A}$在基 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$ 下的方阵表示为 $\mathrm{diag}(A_1,A_2)$．证完．

反之，显然可证明

引理 7.2.4 设$\mathscr{A}$为n维线性空间 $\mathfrak{L}$ 上线性变换．如果在 $\mathfrak{L}$ 中存在一组基$\{\alpha_1,\cdots,\alpha_n\}$，使得$\mathscr{A}$在这组基下的方阵表示为准对角方阵

$$A=\mathrm{diag}(A_1^{(r)},A_2^{(n-r)}),$$

则以$\alpha_1,\cdots,\alpha_r$ 为基之子空间 $\mathfrak{L}_1$ 与以 $\alpha_{r+1},\cdots,\alpha_n$ 为基之子空间 $\mathfrak{L}_2$ 都是线性变换 $\mathscr{A}$ 的不变子空间，且 $\mathfrak{L}$ 有空间直接和分解 $\mathfrak{L}_2=\mathfrak{L}_1\oplus\mathfrak{L}_2$.

这两个引理告诉我们，任给方阵 A，寻找 n 阶非异方阵 P，使得 $P^{-1}AP$ 为准对角形 $\mathrm{diag}(A,\cdots,A_s)$的问题，等价于在 $\mathfrak{L}$ 中寻找 s 个在线性变换$\mathscr{A}$下不变的子空间$\mathfrak{L}_1,\mathfrak{L}_2,\cdots,\mathfrak{L}_s$，使得 $\mathfrak{L}$ 有空间直接和分解：$\mathfrak{L}=\mathfrak{L}_1\oplus\mathfrak{L}_2\oplus\cdots\oplus\mathfrak{L}_s$.

在这一节，我们希望给出方阵在相似下的标准形．为此，我们必须限制在复数范围内讨论．

定义 设 $\mathfrak{L}$ 为 n 维复线性空间，$\mathscr{A}$为 $\mathfrak{L}$ 上线性变换．在 $\mathfrak{L}$ 中取定一组基 $\alpha_1,\cdots,\alpha_n$．记$\mathscr{A}$在这组基下的方阵表示为 A．则 A 的特征多项式，极小多项式和特征根，也分别称为**线性变换$\mathscr{A}$的特征多项式，极小多项式和特征根**.

引理 7.2.5 设 $\mathfrak{L}$ 为 n 维复线性空间，$\mathscr{A}$为 $\mathfrak{L}$ 上线性变换．则$\mathscr{A}$的特征多项式及特征根与基的选取无关.

证 由于不同基下线性变换的方阵表示互相相似．所以只要证明特征多项式和特征根在相似下不改变就行了．事实上，设 P 为任一 n 阶非异方阵，则

$$\det(\lambda I-A)=\det P^{-1}(\lambda I-A)P=\det(\lambda I-P^{-1}AP).$$

于是 A 的特征根和 $P^{-1}AP$ 的特征根也相同．证完.

定义 设$\mathscr{A}$为n维复线性空间$\mathfrak{L}$上线性变换，λ_0为$\mathscr{A}$的特征根. 则$\mathfrak{L}$中非零向量α_0称为**线性变换$\mathscr{A}$的属於特征根λ_0的特征向量**，如果

$$\mathscr{A}(\alpha_0)=\lambda_0\alpha_0.$$

显然有

引理 7.2.6 设$\mathscr{A}$为n维复线性空间$\mathfrak{L}$上线性变换，λ_0为$\mathscr{A}$的特征根，α_0为$\mathfrak{L}$中$\mathscr{A}$的属于特征根λ_0的特征向量，在$\mathfrak{L}$中取定基$\alpha_1,\cdots,\alpha_n$. 记$\mathscr{A}$在这组基下的矩阵表示为A，α_0在这组基下的坐标为x_0，则有

$$Ax_0=\lambda_0x_0,$$

即x_0为n阶方阵A的属于特征根λ_0的特征向量.

将特征向量概念推广，我们有

定义 设$\mathscr{A}$为n维复线性空间$\mathfrak{L}$上线性变换，λ_0为$\mathscr{A}$的特征根. $\mathfrak{L}$中向量α称为线性变换$\mathscr{A}$的属于特征根λ_0的**根向量**，如果存在自然数m，它依赖于α，使得

$$(\mathscr{A}-\lambda_0(id))^m\alpha=0.$$

显然零向量是根向量，又属于特征根λ_0的特征向量也是根向量. 一般有

定理 7.2.1 设$\mathscr{A}$为n维复线性空间$\mathfrak{L}$上线性变换，任取数λ_0，记

$$\mathfrak{L}_{\lambda_0}=\{\alpha\in\mathfrak{L}\mid 存在自然数m=m(\alpha),使得(\mathscr{A}-\lambda_0(id))^m\alpha=0\}.$$

则有

(1) $\mathfrak{L}_{\lambda_0}$为$\mathfrak{L}$的子空间；

(2) $\mathscr{A}(\mathfrak{L}_{\lambda_0})\subset\mathfrak{L}_{\lambda_0}$，即$\mathfrak{L}_{\lambda_0}$为$\mathscr{A}$的不变子空间；

(3) $\mathfrak{L}_{\lambda_0}\neq0$当且仅当$\lambda_0$为$\mathscr{A}$的特征根. 所以当$\lambda_0$为线性变换$\mathscr{A}$的特征根；则$\mathfrak{L}_{\lambda_0}$由所有属于特征根$\lambda_0$的根向量构成，所以$\mathfrak{L}_{\lambda_0}$称为特征根$\lambda_0$的**根子空间**.

证 (1)任取 $\alpha,\beta\in\mathfrak{L}_{\lambda_0}$,于是存在自然数 $m(\alpha),m(\beta)$,使得

$$(\mathscr{A}-\lambda_0(id))^{m(\alpha)}(\alpha)=0,\qquad (\mathscr{A}-\lambda_0(id))^{m(\beta)}(\beta)=0.$$

任取纯量 a,b,记 $m=m(\alpha)+m(\beta)$,于是

$$(\mathscr{A}-\lambda_0(id))^m(a\alpha+b\beta)=a(\mathscr{A}-\lambda_0(id))^m(\alpha)$$
$$+b(\mathscr{A}-\lambda_0(id))^m(\beta)=0.$$

这证明了 $a\alpha+b\beta\in\mathfrak{L}_{\lambda_0}$,即 $\mathfrak{L}_{\lambda_0}$ 为 $\mathfrak{L}$ 的子空间.

(2) 今任取 $\alpha\in\mathfrak{L}_{\lambda_0}$,于是存在自然数 m,使得

$(\mathscr{A}-\lambda_0(id))^m(\alpha)=0.$

因此 $\quad (\mathscr{A}-\lambda_0(id))^m(\mathscr{A}(\alpha))=\mathscr{A}(\mathscr{A}-\lambda_0(id))^m(\alpha)=0.$

这证明了 $\mathscr{A}(\alpha)\in\mathfrak{L}_{\lambda_0}$, $\forall\alpha\in\mathfrak{L}_{\lambda_0}$,所以 $\mathscr{A}(\mathfrak{L}_{\lambda_0})\subset\mathfrak{L}_{\lambda_0}$.

(3) 设 $\mathfrak{L}_{\lambda_0}\neq0$,于是存在非零向量 α 及最小自然数 m,使得 $(\mathscr{A}-\lambda_0(id))^m(\alpha)=0$,自然 $(\mathscr{A}-\lambda_0(id))^{m-1}(\alpha)=\beta\neq0$. 于是 $(\mathscr{A}-\lambda_0(id))\beta=(\mathscr{A}-\lambda_0(id))^m(\alpha)=0$,即有 $\mathscr{A}(\beta)=\lambda_0\beta$. 由 $\beta\neq0$ 便证明了 λ_0 为 $\mathscr{A}$ 的特征根. 反之,若 λ_0 为 $\mathscr{A}$ 之特征根,则有特征向量 $\beta\neq0$,使得 $\mathscr{A}(\beta)=\lambda_0\beta$. 这证明了 $\beta\in\mathfrak{L}_{\lambda_0}$,即 $\mathfrak{L}_{\lambda_0}\neq0$. 证完.

定义 线性空间 $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 称为**幂零线性变换**,如果存在自然数 N,使得 $\mathscr{A}^N=0$. 使得 $\mathscr{A}^N=0$ 成立的最小自然数 N 称为 $\mathscr{A}$ 的**幂零次数**.

显然,零线性变换为幂零次数 1 的幂零线性变换. 又若 $\mathfrak{L}_{\lambda_0}$ 为 $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 关于特征根 λ_0 的根子空间,则 $\mathscr{A}-\lambda_0(id)$ 限制在 $\mathfrak{L}_{\lambda_0}$ 上为幂零线性变换.

显然,幂零线性变换必有零特征根. 反之不一定.

引理 7.2.7 设 $\mathscr{A}$ 为 n 维线性空间 $\mathfrak{L}$ 上线性变换,则 $\mathfrak{L}$ 有不变子空间序列

$$\mathfrak{L}\supseteq\mathscr{A}(\mathfrak{L})\supseteq\mathscr{A}^2(\mathfrak{L})\supseteq\cdots\supseteq\{0\},$$

所以存在自然数 p,使得

$$\mathscr{A}^{p-1}(\mathfrak{L}) \neq \mathscr{A}^{p}(\mathfrak{L}) = \mathscr{A}^{p+1}(\mathfrak{L}) = \cdots = \bigcap_{m=1}^{\infty} \mathscr{A}^{m}(\mathfrak{L})$$

当$\mathscr{A}$为幂零线性变换，则有幂零次数 p; 当$\mathscr{A}$不是幂零线性变换时,则$\mathscr{A}$限制在 $\mathscr{A}^p(\mathfrak{L})$上为非异线性变换

证 由$\mathfrak{L} \supseteq \mathscr{A}(\mathfrak{L})$ 可知 $\mathscr{A}^j(\mathfrak{L}) \supseteq \mathscr{A}^{j+1}(\mathfrak{L}) = \mathscr{A}(\mathscr{A}^j(\mathfrak{L}))$,这证明了$\mathfrak{L} \supseteq \mathscr{A}(\mathfrak{L}) \supseteq \mathscr{A}^2 L \supseteq \cdots$,且$\mathscr{A}^j(\mathfrak{L})$为$\mathscr{A}$的不变子空间,$j=0, 1, 2, \cdots$,这里约定 $\mathscr{A}^0 = id$, 今 $\dim\mathfrak{L} = n$, 所以必存在自然数p使得 $\mathscr{A}^{p-1}(\mathfrak{L}) \neq \mathscr{A}^p(\mathfrak{L}) = \mathscr{A}^{p+1}(\mathfrak{L})$,于是有

$$\mathfrak{L} \supset \mathscr{A}(\mathfrak{L}) \supset \mathscr{A}^2(\mathfrak{L}) \supset \cdots \supset \mathscr{A}^{p-1}(\mathfrak{L}) \neq \mathscr{A}^p(\mathfrak{L}) = \mathscr{A}^{p+1}(\mathfrak{L}) = \cdots$$

于是

$$\bigcap_{m=1}^{\infty} \mathscr{A}^m(\mathfrak{L}) = \mathscr{A}^p(\mathfrak{L}).$$

当$\mathscr{A}$为幂零线性变换，即存在自然数 N 使得 $\mathscr{A}^N = 0$, 于是 $\mathscr{A}^N(\mathfrak{L}) = 0$. 由 $\mathscr{A}^{p-1}(\mathfrak{L}) \neq \mathscr{A}^p(\mathfrak{L}) = \mathscr{A}^{p+1}(\mathfrak{L}) = \cdots$可知 $\mathscr{A}^p(\mathfrak{L}) = 0$,即 $\mathscr{A}^p = 0$, $\mathscr{A}^{p-1} \neq 0$. 所以$\mathscr{A}$的幂零次数为 p.

当$\mathscr{A}$不是幂零线性变换,所以$\mathscr{A}^p(\mathfrak{L}) \neq 0$. 今

$$\mathscr{A}(\mathscr{A}^p(\mathfrak{L})) = \mathscr{A}^{p+1}(\mathfrak{L}) = \mathscr{A}^p(\mathfrak{L}),$$

由定理 7.1.8 可知$\mathscr{A}$为$\mathscr{A}^p(\mathfrak{L})$上非异线性变换. 证完.

引理 7.2.8 设$\mathscr{A}$为n维线性空间$\mathfrak{L}$上线性变换，则$\mathfrak{L}$有不变子空间序列

$$\{0\} \subseteq \ker(\mathscr{A}) \subseteq \ker(\mathscr{A}^2) \subseteq \cdots \subseteq \mathfrak{L}.$$

所以存在自然数q,使得

$$\ker(\mathscr{A}^{q-1}) \neq \ker(\mathscr{A}^q) = \ker(\mathscr{A}^{q+1}) = \cdots = \bigcup_{m=1}^{\infty} \ker(\mathscr{A}^m) = \mathfrak{L}_0.$$

当 $\ker(\mathscr{A}^q) = 0$, 则 $\mathscr{A}$ 为 $\mathfrak{L}$ 上非异线性变换，且 $q=1$; 当 $\ker(\mathscr{A}^q) \neq 0$,则 $\ker(\mathscr{A}^q) = \mathfrak{L}_0$ 为$\mathscr{A}$的属于特征根零的根子空间,所以$\mathscr{A}$限制在 $\mathfrak{L}_0 = \ker(\mathscr{A}^q)$上为幂零线性变换.

证 今任取 $\alpha\in\ker(\mathscr{A}^j)$，则 $\mathscr{A}^j(\alpha)=0$，于是

$$\mathscr{A}^{j+1}(\alpha)=\mathscr{A}^j(\mathscr{A}(\alpha))=\mathscr{A}(\mathscr{A}^j(\alpha))=0.$$

这证明了 $\alpha\in\ker(\mathscr{A}^{j+1})$，且 $\mathscr{A}(\alpha)\in\ker(\mathscr{A}^j)$，即

$$\ker(\mathscr{A}^j)\subseteq\ker(\mathscr{A}^{j+1}),\qquad \mathscr{A}(\ker(\mathscr{A}^j))\subset\ker(\mathscr{A}^j).$$

所以 $\mathfrak{L}\supseteq\cdots\supseteq\ker(\mathscr{A}^2)\supseteq\ker(\mathscr{A})$ 为 $\mathscr{A}$ 之不变子空间序列. 由 $\dim\mathfrak{L}=n$，所以存在自然数 q，使得

$$\ker(\mathscr{A}^{q-1})\neq\ker(\mathscr{A}^q)=\ker(\mathscr{A}^{q+1})=\cdots=\bigcup_{m=1}^{\infty}\ker(\mathscr{A}^m).$$

当 $\ker(\mathscr{A}^q)\neq0$，由根子空间定义可知 $\mathscr{A}$ 的属于特征根零的根子空间

$$\mathfrak{L}_0=\bigcup_{m=1}^{\infty}\ker(\mathscr{A}^m)=\ker(\mathscr{A}^q).$$

当 $\ker(\mathscr{A}^q)=0$，于是 $\mathscr{A}^q$ 非异，因此 $\mathscr{A}$ 非异，所以 $\ker(\mathscr{A})=0$，此即 $q=1$. 证完.

引理 7.2.9（Fitting 引理） 符号同上，记

$$\widetilde{\mathfrak{L}}=\bigcap_{m=1}^{\infty}\mathscr{A}^m(\mathfrak{L}),\qquad \mathfrak{L}_0=\bigcup_{m=1}^{\infty}\ker(\mathscr{A}^m)$$

则 $\mathfrak{L}$ 有 $\mathscr{A}$ 的不变子空间直接和分解

$$\mathfrak{L}=\mathfrak{L}_0\oplus\widetilde{\mathfrak{L}}.$$

当 $\mathfrak{L}_0\neq0$，则 $\mathscr{A}|_{\mathfrak{L}_0}$ 为幂零线性变换；当 $\widetilde{\mathfrak{L}}\neq0$，则 $\mathscr{A}|_{\widetilde{\mathfrak{L}}}$ 为非异线性变换.

证 由引理 7.2.7，若 $\mathscr{A}|_{\widetilde{\mathfrak{L}}}$ 不是非异线性变换，则 $\mathscr{A}^p(\mathfrak{L})=0$，所以 $\widetilde{\mathfrak{L}}=0$；由引理 7.2.8，若 $\mathscr{A}|_{\mathfrak{L}_0}$ 不是幂零线性变换，则 $q=1$，所以 $\mathfrak{L}_0=\bigcup_{m=1}^{\infty}\ker(\mathscr{A}^m)=0$. 这就证明了后一断言. 下面证 $\mathfrak{L}=\mathfrak{L}_0\oplus\widetilde{\mathfrak{L}}=\mathscr{A}^p(\mathfrak{L})\oplus\ker(\mathscr{A}^q)$. 今由定理 7.1.4，

$$\dim\mathfrak{L}=\dim(\mathscr{A}^{p+q}(\mathfrak{L}))+\dim(\ker(\mathscr{A}^{p+q})).$$

由 $\mathscr{A}^{p+q}(\mathfrak{L})=\mathscr{A}^p(\mathfrak{L})$, $\ker(\mathscr{A}^{p+q})=\ker(\mathscr{A}^q)$，所以证明了

$$\dim\mathfrak{L}=\dim\mathfrak{L}_0+\dim\widetilde{\mathfrak{L}}.$$

所以为了证明 $\mathfrak{L}$ 为 $\mathfrak{L}_0$ 和 $\widetilde{\mathfrak{L}}$ 之直接和，只要证 $\mathfrak{L}_0\cap\widetilde{\mathfrak{L}}=0$ 就行了. 当 $\mathfrak{L}_0=0$ 或 $\widetilde{\mathfrak{L}}=0$，则 $\mathfrak{L}_0\cap\widetilde{\mathfrak{L}}=0$. 设 $\mathfrak{L}_0\neq0$, $\widetilde{\mathfrak{L}}\neq0$，任取 $\alpha\in\mathfrak{L}_0\cap\widetilde{\mathfrak{L}}$，由 $\mathfrak{L}_0=\ker(\mathscr{A}^q)$，所以 $\mathscr{A}^q(\alpha)=0$. 但是 $\mathscr{A}^q$ 在 $\widetilde{\mathfrak{L}}$ 上非异，由 $\alpha\in\widetilde{\mathfrak{L}}$，所以 $\alpha=0$. 这证明了 $\mathfrak{L}_0\cap\widetilde{\mathfrak{L}}=0$. 引理证完.

定理 7.2.2 (空间第一分解定理) 设 $\mathscr{A}$ 为 n 维复线性空间 $\mathfrak{L}$ 上线性变换. 记 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 为 $\mathscr{A}$ 的所有不同特征根，则 $\mathfrak{L}$ 有根子空间直接和分解

$$\mathfrak{L}=\mathfrak{L}_{\lambda_1}\oplus\mathfrak{L}_{\lambda_2}\oplus\cdots\oplus\mathfrak{L}_{\lambda_r},$$

且不计次序，分解唯一.

证 取特征根 λ_1，则根子空间 $\mathfrak{L}_{\lambda_1}$ 为线性变换 $\mathscr{A}-\lambda_1(id)$ 的属于特征根零的根子空间. 由 Fitting 引理，$\mathfrak{L}$ 有空间直接和分解

$$\mathfrak{L}=\mathfrak{L}_{\lambda_1}\oplus\widetilde{\mathfrak{L}}_{\lambda_1},$$

其中

$$\widetilde{\mathfrak{L}}_{\lambda_1}=\bigcap_{m=1}^{\infty}(\mathscr{A}-\lambda_1(id))^m(\mathfrak{L}).$$

由于 $\widetilde{\mathfrak{L}}_{\lambda_1}$ 在 $\mathscr{A}-\lambda_1(id)$ 下不变，所以在 $\mathscr{A}$ 下不变. 下面先证 $\mathfrak{L}_{\lambda_2},\cdots,\mathfrak{L}_{\lambda_r}\subset\widetilde{\mathfrak{L}}_{\lambda_1}$. 今 $\mathfrak{L}_{\lambda_j}$ 在 $\mathscr{A}$ 下不变，所以在 $\mathscr{A}-\lambda_1(id)$ 下不变，$j=2,3,\cdots,r$. 先证 $\mathscr{A}-\lambda_1(id)$ 在 $\mathfrak{L}_{\lambda_j}$ 上非异. 设若不然，则存在 $0\neq\alpha\in\mathfrak{L}_{\lambda_j}$, $(\mathscr{A}-\lambda_1(id))(\alpha)=0$，即 $\mathscr{A}(\alpha)=\lambda_1\alpha$. 但是由 $\alpha\in\mathfrak{L}_{\lambda_j}$，所以存在自然数 s，使得 $(\mathscr{A}-\lambda_j(id))^s(\alpha)=0$. 由 $\mathscr{A}(\alpha)=\lambda_1\alpha$，便推出 $(\lambda_1-\lambda_j)^s(\alpha)=0$. 但是 $\lambda_j\neq\lambda_1$，所以 $\alpha=0$. 这导出矛盾. 今 $(\mathscr{A}-\lambda_1(id))|\mathfrak{L}_{\lambda_j}$ 非异，所以在 $\mathfrak{L}_{\lambda_j}$ 上有逆变换 $\mathscr{L}_j$. 任取自然数

m，则 $\mathscr{L}_j^m(\mathfrak{L}_{\lambda_j})=\mathfrak{L}_{\lambda_j}$，所以

$$(\mathscr{A}-\lambda_1(id))^m(\mathfrak{L})\supset(\mathscr{A}-\lambda_1(id))^m\mathfrak{L}_{\lambda_j}$$
$$=(\mathscr{A}-\lambda_1(id))^m\mathscr{L}_j^m(\mathfrak{L}_{\lambda_j})=\mathfrak{L}_{\lambda_j}.$$

这证明了 $\widetilde{\mathfrak{L}}_{\lambda_1}=\bigcap_{m=1}^{\infty}(\mathscr{A}-\lambda_1(id))^m(\mathfrak{L})\supset\mathfrak{L}_{\lambda_j}$, $j=2,3,\cdots,r$. 因此 $\mathscr{A}$ 在 $\widetilde{\mathfrak{L}}_{\lambda_1}$ 上有特征根 $\lambda_2,\lambda_3,\cdots,\lambda_r$,且在 $\widetilde{\mathfrak{L}}_{\lambda_1}$ 中属于特征根 $\lambda_2,\lambda_3,\cdots,\lambda_r$ 之根子空间仍分别为 $\mathfrak{L}_{\lambda_2},\cdots,\mathfrak{L}_{\lambda_r}$. 再 $\mathscr{A}$ 在 $\widetilde{\mathfrak{L}}_{\lambda_1}$ 上没有特征根 λ_1，否则有属于特征根 λ_1 之特征向量 $\alpha_1\in\widetilde{\mathfrak{L}}_{\lambda_1}$，即 $\widetilde{\mathfrak{L}}_{\lambda_1}\cap\mathfrak{L}_{\lambda_1}\neq 0$. 这和 Fitting 引理矛盾.

由归纳法假设，$\widetilde{\mathfrak{L}}_{\lambda_1}$ 分解为根子空间直接和

$$\widetilde{\mathfrak{L}}_{\lambda_1}=\mathfrak{L}_{\lambda_2}\oplus\cdots\oplus\mathfrak{L}_{\lambda_r}.$$

于是

$$\mathfrak{L}=\mathfrak{L}_{\lambda_1}\oplus\widetilde{\mathfrak{L}}_{\lambda_1}=\mathfrak{L}_{\lambda_1}\oplus\mathfrak{L}_{\lambda_2}\oplus\cdots\oplus\mathfrak{L}_{\lambda_r}.$$

分解唯一性的证明，由上面的构造过程便知. 定理证完.

定理 7.2.3 设 A 为 n 阶复方阵，$\lambda_1,\lambda_2,\cdots,\lambda_r$ 为 A 的所有不同特征根. 则 A 相似于准对角阵

$$\mathrm{diag}(A_1,A_2\cdots,A_r),$$

其中 $A_j-\lambda_jI$ 为幂零方阵，所以 A_j 恰有一个特征根 λ_j，$j=1,2,\cdots,r$.

证 设 $\mathfrak{L}$ 为 n 维复线性空间，$\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 $\mathfrak{L}$ 的一组基. 显然在 $\mathfrak{L}$ 上存在线性变换 $\mathscr{A}$，使得 $\mathscr{A}$ 关于基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的方阵表示为 A. 由空间第一分解定理，$\mathfrak{L}$ 有关于 $\mathscr{A}$ 之根子空间直接和分解 $\mathfrak{L}=\mathfrak{L}_{\lambda_1}\oplus\cdots\oplus\mathfrak{L}_{\lambda_r}$. 在每个 $\mathfrak{L}_{\lambda_j}$ 中取基，拼成 $\mathfrak{L}$ 之基. 于是证明了 A 相似于准对角方阵 $\mathrm{diag}(A_1,A_2,\cdots,A_r)$，其中 $\mathscr{A}|_{\mathfrak{L}_{\lambda_j}}$ 之方阵表示为 A_j, $j=1,2,\cdots,r$. 于是 $A_j-\lambda_j(I)$ 为幂零方阵，且 A_j 恰

有一个特征根 λ_j. 定理证完.

为了进一步将 n 阶复方阵 A 在相似下化简，问题化为如何进一步在相似下化简复方阵 $A_1, A_2, \cdots, A_r$. 换句话说，只要考虑线性空间 $\mathfrak{L}$ 上线性变换 $\mathscr{A}$，使得 $\mathscr{A}$ 在 $\mathfrak{L}$ 上恰有一个特征根 λ_0，所以 $\mathscr{A}-\lambda_0(id)$ 在 $\mathfrak{L}$ 上幂零. 即化为考虑在幂零线性变换作用下，如何将线性空间分解为若干个不变子空间的直接和. 为此，先引进一般的循环子空间的概念.

定义 设 $\mathscr{A}$ 为 n 维线性空间 $\mathfrak{L}$ 上线性变换. 任取 $\mathfrak{L}$ 上非零向量 α, 则由 $\mathfrak{L}$ 的子集合 $\{\alpha, \mathscr{A}(\alpha), \mathscr{A}^2(\alpha), \cdots\}$ 线性生成的子空间称为 $\mathscr{A}$ 的由 α 生成的**循环子空间**, α 称为**循环向量**.

引理 7.2.10 设 $\mathscr{A}$ 为 n 维线性空间 $\mathfrak{L}$ 上线性变换，则 $\mathscr{A}$ 的由 $\alpha\neq 0$ 生成的循环子空间 $\mathfrak{R}(\alpha)$ 必为 $\mathscr{A}$ 的不变子空间. 且存在自然数 N, 使得 $\alpha, \mathscr{A}(\alpha), \mathscr{A}^2(\alpha), \cdots, \mathscr{A}^{N-1}(\alpha)$ 为它的一组基.

证 今 $\mathscr{A}(\{\alpha, \mathscr{A}(\alpha), \cdots\}) = \{\mathscr{A}(\alpha), \mathscr{A}^2(\alpha), \cdots\}$，所以 $\mathfrak{R}(\alpha)$ 为 $\mathscr{A}$ 的不变子空间. 由于 $\dim\mathfrak{R}(\alpha)\leqslant\dim\mathfrak{L}=n$，所以对 $\alpha, \mathscr{A}(\alpha)$, $\mathscr{A}^2(\alpha), \cdots$ 必存在自然数 N, 使得 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{N-1}(\alpha)$ 线性无关，而 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{N-1}(\alpha), \mathscr{A}^N(\alpha)$ 线性相关，即有

$$\mathscr{A}^N(\alpha) = \sum_{j=0}^{N-1} a_j \mathscr{A}^j(\alpha).$$

由归纳法不难证明 $\mathscr{A}^{N+1}(\alpha), \mathscr{A}^{N+2}(\alpha), \cdots$ 都是 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{N-1}(\alpha)$ 的线性组合. 这证明了 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{N-1}(\alpha)$ 为 $\mathfrak{R}(\alpha)$ 的基. 引理证完.

引理 7.2.11 设 $\mathscr{A}$ 为 n 维线性空间 $\mathfrak{L}$ 上幂零线性变换. 则 $\mathscr{A}$ 的由 $\alpha\neq 0$ 生成的循环子空间 $\mathfrak{R}(\alpha)$ 有基 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{N-1}(\alpha)$ 且 $\mathscr{A}^N(\alpha)=0$.

证 今 $\mathscr{A}$ 为幂零线性变换，所以存在最小自然数 N 使得 $\mathscr{A}^N(\alpha)=0$, $\mathscr{A}^{N-1}(\alpha)\neq 0$. 下面证明 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{N-1}(\alpha)$ 线性

无关. 设若不然，则存在不全为零的数 c_0, c_1, $\cdots$, c_{N-1}, 使得 $c_0\alpha+c_1\mathscr{A}(\alpha)+\cdots+c_{N-1}\mathscr{A}^{N-1}(\alpha)=0$. 设 $c_0=\cdots=c_{r-1}=0, c_r\neq 0$. 则 $0=\mathscr{A}^{N-1-r}(0)=\mathscr{A}^{N-1-r}(c_r\mathscr{A}^r(\alpha)+c_{r+1}\mathscr{A}^{r+1}(\alpha)+\cdots+c_{N-1}\mathscr{A}^{N-1}(\alpha))=c_r\mathscr{A}^{N-1}(\alpha)+c_{r+1}\mathscr{A}^N(\alpha)+\cdots+c_{N-1}\mathscr{A}^{2N-2-r}(\alpha)=c_r\mathscr{A}^{N-1}(\alpha)$. 这导出 $c_r=0$, 所以推出矛盾. 因此证明了 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{N-1}(\alpha)$ 为 $\mathscr{A}$ 的由 α 生成的循环子空间 $\mathfrak{N}(\alpha)$ 的基. 证完.

定理 7.2.4 (空间第二分解定理) 设 $\mathscr{A}$ 为 n 维线性空间 $\mathfrak{L}$ 上的幂零线性变换，则 $\mathfrak{L}$ 可分解为循环子空间直接和，且如不计次序，则在线性同构意义下分解唯一.

证 对 $\mathfrak{L}$ 的维数作归纳法. 当 $n=1$ 时定理显然成立. 假设对维数小于 n 的线性空间定理成立. 现在考虑 n 维线性空间 $\mathfrak{L}$. 今 $\mathscr{A}$ 为幂零线性变换. 由引理 7.1.8, 所以 $\dim\mathscr{A}(\mathfrak{L})<n$. 由于 $\mathscr{A}$ 限制在 $\mathscr{A}(\mathfrak{L})$ 上仍为幂零线性变换，由归纳法假设，$\mathscr{A}(\mathfrak{L})$ 分解为循环子空间直接和. 由引理 7.2.11, 所以 $\mathscr{A}(\mathfrak{L})$ 有基

$$\begin{array}{cccc}\mathscr{A}^{N_1}(\alpha_1), & \mathscr{A}^{N_1-1}(\alpha_1), & \cdots, & \mathscr{A}(\alpha_1),\\ \vdots & \vdots & & \vdots\\ \mathscr{A}^{N_s}(\alpha_s), & \mathscr{A}^{N_s-1}(\alpha_s), & \cdots, & \mathscr{A}(\alpha_s),\end{array}$$

其中 $\alpha_1,\alpha_2,\cdots,\alpha_s\in\mathfrak{L}$, N_j 为自然数，适合 $\mathscr{A}^{N_j}(\alpha_j)\neq 0, \mathscr{A}^{N_j+1}(\alpha_j)=0, 1\leqslant j\leqslant s$.

先证 $\mathscr{A}^{N_1}(\alpha_1),\cdots,\mathscr{A}^{N_s}(\alpha_s)$ 为 $\mathscr{A}(\mathfrak{L})\cap\ker(\mathscr{A})$ 之基. 事实上，任取 $\beta\in\mathscr{A}(\mathfrak{L})\cap\ker(\mathscr{A})$, 则 $\beta=\Sigma a_{ij}\mathscr{A}^i(\alpha_j)$. 由 $\mathscr{A}(\beta)=0$ 有 $\Sigma a_{ij}\mathscr{A}^{i+1}(\alpha_j)=0$, 因此 $a_{ij}=0$ 当 $i+1\leqslant N_j$ 时成立. 所以 $\beta=\alpha_{N_11}\mathscr{A}^{N_1}(\alpha_1)+\cdots+a_{N_ss}\mathscr{A}^{N_s}(\alpha_s)$. 这证明了断言.

在 $\ker(\mathscr{A})$ 中选取向量 $\alpha_{s+1}, \cdots, \alpha_t$ 使得 $\mathscr{A}^{N_1}(\alpha_1)\cdots, \mathscr{A}^{N_s}(\alpha_s), \alpha_{s+1},\cdots,\alpha_t$ 为 $\ker(\mathscr{A})$ 之基. 于是 $\mathfrak{L}$ 有向量组 $\mathfrak{S}$

$$\begin{array}{l}\mathscr{A}^{N_1}(\alpha_1),\ \mathscr{A}^{N_1-1}(\alpha_1),\ \cdots,\ \mathscr{A}(\alpha_1),\ \alpha_1,\\ \cdots\cdots\\ \mathscr{A}^{N_s}(\alpha_s),\ \mathscr{A}^{N_s-1}(\alpha_s),\ \cdots,\ \mathscr{A}(\alpha_s),\ \alpha_s,\end{array}$$

$$\alpha_{s+1},$$
$$\vdots$$
$$\alpha_t.$$

下面证明 $\mathfrak{S}$ 构成 $\mathfrak{V}$ 的一组基. 事实上，任取 $\alpha\in\mathfrak{V}$，则 $\mathscr{A}(\alpha)\in\mathscr{A}(\mathfrak{V})$. 由归纳法假设，

$$\mathscr{A}(\alpha)=\sum_{i=1}^{s}\sum_{j=1}^{N_i}a_{i,j-1}\mathscr{A}^{j}(\alpha_i).$$

令

$$\beta=\alpha-\sum_{i=1}^{s}\sum_{j=1}^{N_i}a_{i,j-1}\mathscr{A}^{j-1}(\alpha_i)$$

则 $\mathscr{A}(\beta)=0$，所以 $\beta\in\ker(\mathscr{A})$，因此

$$\beta=\sum_{i=1}^{s}a_{iN_i}\mathscr{A}^{N_i}(\alpha_i)+\sum_{i=s+1}^{t}a_i\alpha_i,$$

代回有

$$\begin{aligned}\alpha&=\beta+\sum_{i=1}^{s}\sum_{j=1}^{N_i}a_{i,j-1}\mathscr{A}^{j-1}(\alpha_i)\\&=\sum_{i=1}^{s}\sum_{j=0}^{N_i}a_{ij}\mathscr{A}^{j}(\alpha_i)+\sum_{i=s+1}^{t}a_i\alpha_i.\end{aligned}$$

这证明了 $\mathfrak{V}$ 由向量集 $\mathfrak{S}$ 线性生成. 所以为了证 $\mathfrak{S}$ 为 $\mathfrak{V}$ 的一组基，只要证它们线性无关即可. 今若

$$\sum_{i=1}^{s}\sum_{j=0}^{N_i}a_{ij}\mathscr{A}^{j}(\alpha_i)+\sum_{i=s+1}^{t}a_i\alpha_i=0.$$

作用 $\mathscr{A}$，有

$$\sum_{i=1}^{s}\sum_{j=0}^{N_i-1}a_{ij}\mathscr{A}^{j+1}(\alpha_i)=0.$$

由归纳法假设，$a_{ij}=0$，$0\leqslant j<N_i$，$1\leqslant i\leqslant s$. 于是

$$\sum_{i=1}^{s}a_{iN_i}\mathscr{A}^{N_i}(\alpha_i)+\sum_{i=s+1}^{t}a_i\alpha_i=0.$$

但是 $\mathscr{A}^{N_1}(\alpha_1),\cdots,\mathscr{A}^{N_s}(\alpha_s),\alpha_{s+1},\cdots,\alpha_t$ 为 $\ker(\mathscr{A})$ 之基，所以 $a_{ij}=0,\ 0\leqslant j\leqslant N_i,\ 1\leqslant i\leqslant s,\ a_i=0,s+1\leqslant i\leqslant t$. 这证明了向量组 $\mathfrak{S}$ 线性无关，即为 $\mathfrak{L}$ 的基.

记 $\alpha_1,\cdots,\alpha_t$ 生成 $\mathscr{A}$ 的循环子空间分别为 $\mathfrak{L}_1,\cdots,\mathfrak{L}_t$. 由引理 7.2.11 可知 $\alpha_j,\mathscr{A}(\alpha_j),\cdots,\mathscr{A}^{N_j}(\alpha_j)$ 为 $\mathfrak{L}_j$ 之基，$1\leqslant j\leqslant s$，α_k 为 $\mathfrak{L}_k$ 之基 $s+1\leqslant k\leqslant t$. 因此也证明了 $\mathfrak{L}$ 有循环子空间直接和分解

$$\mathfrak{L}=\mathfrak{L}_1\oplus\cdots\oplus\mathfrak{L}_t.$$

余下证上述分解在同构意义下如不计次序则唯一. 设

$$\begin{aligned}\mathfrak{L}&=\mathfrak{L}_1\oplus\cdots\oplus\mathfrak{L}_s\oplus\mathfrak{L}_{s+1}\oplus\cdots\oplus\mathfrak{L}_t\\&=\mathfrak{M}_1\oplus\cdots\oplus\mathfrak{M}_p\oplus\mathfrak{M}_{p+1}\oplus\cdots\oplus\mathfrak{M}_q,\end{aligned}$$

其中 $\quad\dim(\mathfrak{L}_i)>1,\ 1\leqslant i\leqslant s,\ \dim(\mathfrak{L}_j)=1,\ s+1\leqslant j\leqslant t;$

$\dim(\mathfrak{M}_i)>1,1\leqslant i\leqslant p,\ \dim(\mathfrak{M}_j)=1,p+1\leqslant j\leqslant q.$

于是 $\quad\mathscr{A}(\mathfrak{L}_1)\oplus\cdots\oplus\mathscr{A}(\mathfrak{L}_s)=\mathscr{A}(\mathfrak{M}_1)\oplus\cdots\oplus\mathscr{A}(\mathfrak{M}_p).$

由归纳法假设，有 $p=s$，且若不计次序，无妨设

$$\dim(\mathscr{A}(\mathfrak{L}_i))=\dim(\mathscr{A}(\mathfrak{M}_i)),\qquad i=1,2,\cdots,s.$$

今 $\dim(\mathscr{A}(\mathfrak{L}_i))=\dim(\mathfrak{L}_i)-1,\dim(\mathscr{A}(\mathfrak{M}_i))=\dim(\mathfrak{M}_i)-1$，所以证明了 $\dim(\mathfrak{L}_i)=\dim(\mathfrak{M}_i),1\leqslant i\leqslant s=p$. 另一方面，

$$\dim(\mathfrak{L})=\sum_{j=1}^{t}\dim(\mathfrak{L}_j)=\sum_{j=1}^{s}\dim(\mathfrak{L}_j)+(t-s),$$

$$\begin{aligned}\dim(\mathfrak{L})&=\sum_{k=1}^{q}\dim(\mathfrak{M}_k)=\sum_{k=1}^{s}\dim(\mathfrak{M}_k)+(q-s)\\&=\sum_{k=1}^{s}\dim(\mathfrak{L}_k)+(q-s).\end{aligned}$$

这证明了 $q=t$；且 $\dim(\mathfrak{L}_j)=\dim(\mathfrak{M}_j)$，所以 $\mathfrak{L}_j$ 和 $\mathfrak{M}_j$ 线性同构，$1\leqslant j\leqslant t$. 定理证完.

将上面两个空间分解定理合在一起，注意到 $\mathscr{A}-\lambda_i(id)$ 在根子空间 $\mathfrak{L}_{\lambda_i}$ 上为幂零线性变换，所以证明了

定理 7.2.5 设$\mathscr{A}$为n维复线性空间$\mathfrak{L}$上线性变换，λ_1，λ_2，$\cdots,\lambda_r$为$\mathscr{A}$的全部不同特征根，则$\mathfrak{L}$有根子空间直接和分解

$$\mathfrak{L}=\mathfrak{L}_{\lambda_1}\oplus\cdots\oplus\mathfrak{L}_{\lambda_r},$$

分解不计次序则唯一．而每个根子空间$\mathfrak{L}_{\lambda_j}$又可分解为关于幂零线性变换$\mathscr{A}-\lambda_j(id)$的循环子空间直接和

$$\mathfrak{L}_{\lambda_j}=\mathfrak{L}_{j1}\oplus\mathfrak{L}_{j2}\oplus\cdots\oplus\mathfrak{L}_{js_j},\qquad j=1,2,\cdots,r.$$

且此分解若不计次序则在同构意义下唯一．

最后，给出上面定理的矩阵形式．即给出复方阵在相似下的 Jordan 标准形．为此先给出

引理 7.2.12 设$\mathscr{A}$为n维线性空间$\mathfrak{L}$上幂零线性变换，且$\mathfrak{L}$为$\mathscr{A}$的由$\alpha\neq 0$生成的循环空间．则在$\mathfrak{L}$的基$\mathscr{A}^{n-1}(\alpha),\mathscr{A}^{n-2}(\alpha),\cdots,\mathscr{A}(\alpha),\alpha$下，线性变换$\mathscr{A}$的方阵表示为

$$N^{(n)}=\begin{pmatrix}0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1\\ & & & 0\end{pmatrix}.$$

引理 7.2.13 设$\mathscr{A}$为n维复线性空间$\mathfrak{L}$上线性变换，且存在复数λ_0使得$\mathscr{A}-\lambda_0(id)$为$\mathfrak{L}$上幂零线性变换．设$\mathfrak{L}$为$\mathscr{A}-\lambda_0(id)$的由$\alpha\neq 0$生成的循环空间，则在$\mathfrak{L}$的基

$$(\mathscr{A}-\lambda_0(id))^{n-1}(\alpha),\cdots,(\mathscr{A}-\lambda_0(id))(\alpha),\quad \alpha$$

下线性变换$\mathscr{A}$的方阵表示为

$$J_{\lambda_0}^{(n)}=\lambda_0 I^{(n)}+N^{(n)}=\begin{pmatrix}\lambda_0 & 1 & & & \\ & \lambda_0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_0 & 1\\ & & & & \lambda_0\end{pmatrix}.$$

这个方阵称为属于特征根λ_0的 Jordan 块．

由引理 7.2.12，引理 7.2.13 以及定理 7.2.5 便证明了

定理 7.2.6(复方阵的 Jordan 标准形定理) n 阶复方阵 A 必相似于 Jordan 标准形

$$\mathrm{diag}(J_1, J_2, \cdots, J_t)$$

其中 $J_i=\lambda_i I^{(m_i)}+N^{(m_i)}$ 为 m_i 阶 Jordan 块，$i=1, 2, \cdots, t$.（这里 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 并不要求两两不等）. 又不计对角块的次序，则 Jordan 标准形唯一. 又 Jordan 块 $J_1, \cdots, J_t$ 之特征多项式 $(\lambda-\lambda_1)^{m_1}, (\lambda-\lambda_2)^{m_2}, \cdots, (\lambda-\lambda_s)^{m_s}$ 称为 A 的**初等因式**.

关于如何计算 n 阶复方阵的 Jordan 标准形，即如何计算它的初等因式组. 这可以在第十三章中找到.

习题 7.2

1. n 维线性空间 $\mathfrak{L}$ 为其上线性变换 $\mathscr{A}$ 的循环空间当且仅当对 $\mathscr{A}$ 的任一特征根决定的特征向量全体线性生成一维子空间；当且仅当 $\mathscr{A}$ 的特征多项式等于 $\mathscr{A}$ 的**极小多项式**（它们分别定义为 $\mathscr{A}$ 的任一方阵表示的特征多项式和极小多项式）.

2. 试证：n 维线性空间 $\mathfrak{L}$ 上线性变换若有一个方阵表示为对角方阵，则 $\mathscr{A}$ 有 n 个线性无关的特征向量. 反之亦然.

3. 试证：(i) 复线性空间 $\mathfrak{L}$ 上一组两两可交换的线性变换有公共特征向量. (ii) 用 (i) 证明：一组两两可交换的 n 阶复方阵能用一个公共的 n 阶非异复方阵，将它们同时相似于上三角方阵，(iii) 特别，当 (ii) 中给出的那组两两可交换的 n 阶方阵的 Jordan 标准形都是对角方阵的情形，则它们同时相似于对角形.

4. 设 A 和 B 都是 n 阶方阵，且 B 幂零，又 $AB=BA$. 试证：

$$\det(\lambda I-(A+B))=\det(\lambda I-A).$$

5. 设 $\lambda_1, \cdots, \lambda_k$ 为 n 阶方阵 A 的 k 个不同特征根，试证：

$$n(k-1)\leqslant\sum_{j=1}^{k}\mathrm{rank}(\lambda_j I-A).$$

6. 设 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 为 n 维线性空间 $\mathfrak{L}$ 中关于线性变换 $\mathscr{A}$ 的不变子空间，什么时候 $\mathscr{A}(\mathfrak{L}_1\cap\mathfrak{L}_2)=\mathscr{A}(\mathfrak{L}_1)\cap\mathscr{A}(\mathfrak{L}_2)$？

7. 设 n 维线性空间 $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 的特征多项式为 $f(\lambda)$. 设 $f(\lambda)=$

$g(\lambda)h(\lambda)$，其中$(g(\lambda),h(\lambda))=1$．试证：$\mathfrak{L}$中存在$\mathscr{A}$的不变子空间$\mathfrak{L}_1$和$\mathfrak{L}_2$，使得$g(\mathscr{A}|\mathfrak{L}_1)=0, h(\mathscr{A}|\mathfrak{L}_2)=0$，又$\mathfrak{L}=\mathfrak{L}_1\oplus\mathfrak{L}_2$为子空间直接和．用这一事实，试给出第一空间分解定理的新证明．

8．设$\alpha_1,\cdots,\alpha_n$为n维线性空间$\mathfrak{L}$中一组基，且它们都是$\mathfrak{L}$上线性变换$\mathscr{A}$的特征向量，又$\mathscr{A}$的特征根都不相同．试证：$\mathfrak{L}$为$\mathscr{A}$的循环空间，并找出循环向量．

9．设$\mathscr{A}$为n维线性空间$\mathfrak{L}$上线性变换，λ_j为$\mathscr{A}$的e_j重特征根．试证：关于特征根λ_j的根子空间$\mathfrak{L}_{\lambda_j}$可定义为$\mathfrak{L}_{\lambda_j}=\{\alpha\in\mathfrak{L}(\mathscr{A}-\lambda_j(id))^{e_j}(\alpha)=0\}$．

10．试证：两两可交换的$2n+1$阶实方阵必有公共的实特征向量．

11．设$\mathscr{A}$为n维线性空间$\mathfrak{L}$上线性变换，$f(x)$，$g(x)$为普通的多项式．试证：

$$\ker(f(\mathscr{A}))+\ker(g(\mathscr{A}))\subset\ker(f(\mathscr{A}),g(\mathscr{A})).$$

又设当$(f,g)=1$，则有下面子空间直接和分解

$$\ker(f(\mathscr{A}))\oplus\ker(g(\mathscr{A}))=\ker(f(\mathscr{A}),g(\mathscr{A})).$$

12．试证：线性空间$\mathfrak{L}$上线性变换$\mathscr{A}$和$\mathscr{B}$若有$\mathscr{A}^2=id$，$\mathscr{B}^2=id$，$\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$，则在$\mathfrak{L}$中存在一组基，使得$\mathscr{A}$和$\mathscr{B}$的方阵表示都是对角方阵，且对角元素为1或$-1$．

13．设$0=\mathfrak{L}_0,\mathfrak{L}_1,\cdots,\mathfrak{L}_m,\mathfrak{L}_{m+1}=0$都是线性空间．设存在线性映射$\mathscr{A}_i$：$\mathfrak{L}_i\longrightarrow\mathfrak{L}_{i+1}$，$i=0,1,\cdots,m$，它们有

$$\ker(\mathscr{A}_{i+1})=\mathscr{A}_i(\mathscr{A}_i),\quad i=0,1,\cdots,m-1.$$

试证：

$$\sum_{i=0}^{m}(-1)^i\dim\mathfrak{L}_i=0.$$

14．(Schur 引理) n维线性空间$\mathfrak{L}$上线性变换集$\mathfrak{S}$称为**不可约的**，如果除了$\mathfrak{L}$及零空间外，没有其他在$\mathfrak{S}$中所有线性变换作用下都不变的子空间．

设$\mathfrak{S}_1$和$\mathfrak{S}_2$分别为线性空间$\mathfrak{L}_1$和$\mathfrak{L}_2$上的不可约线性变换集．设存在$\mathfrak{L}_1$到$\mathfrak{L}_2$内的线性映射σ，使得任取$\mathscr{A}\in\mathfrak{S}_1$，则存在$\mathscr{B}\in\mathfrak{S}_2$，使得$\sigma\circ\mathscr{A}=\mathscr{B}\circ\sigma$．且当$\mathscr{A}$遍历$\mathfrak{S}_1$时，$\mathscr{B}$也遍历$\mathfrak{S}_2$．试证：$\sigma=0$或者$\dim\mathfrak{L}_1=\dim\mathfrak{L}_2$，且$\sigma$为$\mathfrak{L}_1$到$\mathfrak{L}_2$上的线性同构．

特别，当$\mathfrak{L}_1=\mathfrak{L}_2$且$\mathfrak{L}_1$为复线性空间，又当$\mathfrak{S}_1=\mathfrak{S}_2$，$\sigma\circ\mathscr{A}=\mathscr{A}\circ\sigma$，$\forall\mathscr{A}\in$

$\mathfrak{S}_1$,则 σ 为 $\mathfrak{L}_1$ 上的纯量线性变换,即存在复数 a,使得 $\sigma=a(id)$.

15. 试不用 Jordan 标准形理论,直接证明: n 阶幂等方阵 A(即适合 $A^2=A$)相似于对角形 $\mathrm{diag}(I^{(r)},0^{(n-r)})$,其中 $r=\mathrm{rank}(A)$. 用这一结果证明: n 阶幂等方阵 A 有

$$\mathrm{tr}A=\mathrm{rank}(A).$$

16. 设 $A_1,\cdots,A_m$ 为 n 阶方阵,适合条件

$$A_1+A_2+\cdots+A_m=I^{(n)}$$

试证下面三条件互相等价.

(i) $A_1,A_2,\cdots,A_m$ 都幂等;

(ii) $\sum_{j=1}^{m}\mathrm{rank}(A_j)=n$;

(iii) $A_iA_j=0,\ i\neq j$.

且这时存在 n 阶非异方阵 P,使得

$$PA_jP^{-1}=\mathrm{diag}(0^{(n_1)},\cdots,0^{(n_{j-1})},I^{(n_j)},0^{(n_{j+1})},\cdots,0^{(n_m)}),$$

$$1\leqslant j\leqslant m.$$

第八章　多重线性函数

§8.1　线性函数和双线性函数

定义　给定n维复(实、有理)数域上线性空间$\mathfrak{L}$. 自变量取$\mathfrak{L}$中向量，函数值为复(实、有理)数的函数f称为**线性函数**，如果

$f(a\alpha+b\beta)=af(\alpha)+bf(\beta)$，$\forall\alpha$，$\beta\in\mathfrak{L}$，$a$，$b$为复(实、有理)数.

定义　设f，g为n维线性空间$\mathfrak{L}$上线性函数，则$\alpha\to f(\alpha)+g(\alpha)$，$\forall\alpha\in\mathfrak{L}$定义了$\mathfrak{L}$上线性函数，记作$f+g$，称为$f$和$g$之**和**. 给定数$a$，则$\alpha\to af(\alpha)$，$\forall\alpha\in\mathfrak{L}$定义了$\mathfrak{L}$上线性函数，称为$f$关于纯量$a$的**纯量乘积**.

于是有

定理8.1.1　记$\mathfrak{L}^*$为n维线性空间$\mathfrak{L}$上所有线性函数构成的集合，则$\mathfrak{L}^*$在上面定义的加法及纯量乘积下仍构成线性空间. 称为线性空间$\mathfrak{L}$的**共轭空间**或称**对偶空间**.

下面给定n维线性空间$\mathfrak{L}$，我们来具体描述共轭空间，即给出线性函数的坐标表达式.

在$\mathfrak{L}$中取定基$\alpha_1,\cdots,\alpha_n$，于是$\mathfrak{L}$中任一向量

$$\alpha=\sum_{i=1}^{n}x_i\alpha_i,$$

其中x_1，x_2，$\cdots$，x_n为n个独立自变量. 设f为$\mathfrak{L}$上线性函数，于是

$$f(\alpha)=f\Big(\sum_{j=1}^{n}x_j\alpha_j\Big)=\sum_{j=1}^{n}x_jf(\alpha_j)=\sum_{j=1}^{n}a_jx_j,$$

其中 $a_j=f(\alpha_j)$，$j=1,2,\cdots,n$. 反之，任取常数 $a_1,\cdots,a_n$，则

$$f\left(\sum_{j=1}^{n} x_j\alpha_j\right)=\sum_{j=1}^{n}a_jx_j$$

为线性函数，它有 $a_j=f(\alpha_j)$，$j=1,2,\cdots,n$. 所以记

$$x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\qquad f=\begin{pmatrix}f(\alpha_1)\\f(\alpha_2)\\\vdots\\f(\alpha_n)\end{pmatrix},$$

则线性函数 $f(\alpha)$ 和 $n\times 1$ 矩阵 f 之间有自然的一一对应关系. 且若

$$f(\alpha)\to\begin{pmatrix}f(\alpha_1)\\f(\alpha_2)\\\vdots\\f(\alpha_n)\end{pmatrix},\qquad g(\alpha)\to\begin{pmatrix}g(\alpha_1)\\g(\alpha_2)\\\vdots\\g(\alpha_n)\end{pmatrix},$$

又任取常数 a,b，则有

$$(af+bg)(\alpha)=af(\alpha)+bg(\alpha)\to af+bg$$

$$=\begin{pmatrix}(af+bg)(\alpha_1)\\(af+bg)(\alpha_2)\\\vdots\\(af+bg)(\alpha_n)\end{pmatrix},$$

这证明了共轭空间 $\mathfrak{L}^*$ 和 $n\times 1$ 矩阵构成的线性空间 V_n 间在对应 $f(\alpha)\to f$ 下为线性同构，所以 $\dim\mathfrak{L}^*=n$.

定义 给定 n 维线性空间 $\mathfrak{L}$，记 $\mathfrak{L}^*$ 为 $\mathfrak{L}$ 的共轭空间. 在 $\mathfrak{L}$ 中任取基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，则在 $\mathfrak{L}^*$ 中存在基 $f_1,f_2,\cdots,f_n$，使得

$$f_i(\alpha_j)=\delta_{ij},\qquad i,j=1,2,\cdots,n.$$

基 $f_1,f_2,\cdots,f_n$ 称为 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的**对偶基**.

定理 8.1.2 设 $\mathfrak{L}$ 为 n 维线性空间，$\mathfrak{L}^*$ 为 $\mathfrak{L}$ 的对偶空间. 在

$\mathfrak{L}$ 中任取基 $\alpha_1, \alpha_2, \cdots, \alpha_n$，$\mathfrak{L}^*$ 中对偶基记作 $f_1, f_2, \cdots, f_n$. 则任取 $f \in \mathfrak{L}^*$，有

$$f = \sum_{i=1}^{n} f(\alpha_i) f_i.$$

证 任取 $\alpha = \sum_{j=1}^{n} x_j \alpha_j$，则

$$\left(f - \sum_{i=1}^{n} f(\alpha_i) f_i\right)(\alpha) = f(\alpha) - \sum_{i=1}^{n} f(\alpha_i) f_i(\alpha)$$

$$= \sum_{j=1}^{n} x_j f(\alpha_j) - \sum_{i,j=1}^{n} f(\alpha_i) x_j f_i(\alpha_j)$$

$$= \sum_{j=1}^{n} f(\alpha_j) x_j - \sum_{i,j=1}^{n} \delta_{ij} x_j f(\alpha_i) = 0.$$

这证明了 $f = \sum_{i=1}^{n} f(\alpha_i) f_i$. 定理证完.

定理 8.1.3 在 n 维线性空间 $\mathfrak{L}$ 中取定基 $\{\alpha_1, \cdots, \alpha_n\}$ 和 $\{\beta_1, \beta_2, \cdots, \beta_n\}$，相应对偶空间 $\mathfrak{L}^*$ 中对偶基为 $\{f_1, \cdots, f_n\}$ 和 $\{g_1, \cdots, g_n\}$. 若 $\mathfrak{L}$ 的基变换对应的非异方阵为 $A = (a_{ij})$，则对偶基的基变换对应的非异方阵为 $(A')^{-1}$.

证 令

$$\beta_i = \sum_{j=1}^{n} a_{ji} \alpha_j, \qquad i = 1, 2, \cdots, n.$$

记

$$g_i = \sum_{j=1}^{n} b_{ji} f_j, \qquad i = 1, 2, \cdots, n.$$

而

$$\delta_{ij} = g_i(\beta_j) = \sum_{l=1}^{n} b_{li} f_l\left(\sum_{k=1}^{n} a_{kj} \alpha_k\right)$$

$$=\sum_{l,k=1}^{n} b_{li}a_{kj}f_l(\alpha_k) = \sum_{l=1}^{n} b_{ki}a_{kj}.$$

这证明了 n 阶方阵 $B=(b_{ij})$ 有 $B'A=I$，所以 $B=(A^{-1})'$. 证完.

定理 8.1.4 设$\mathscr{A}$为 n 维线性空间 $\mathfrak{L}$ 上线性变换. 任取 $\mathfrak{L}$ 上线性函数 f，则

$$\alpha \to f(\mathscr{A}(\alpha)), \qquad \forall \alpha \in \mathfrak{L}$$

为 $\mathfrak{L}$ 上线性函数，记作 $\mathscr{A}^*(f)$，即有

$$(\mathscr{A}^*(f))(\alpha)=f(\mathscr{A}(\alpha)), \qquad \forall \alpha \in \mathfrak{L}.$$

于是引进了 $\mathfrak{L}$ 的对偶空间 $\mathfrak{L}^*$ 上映射$\mathscr{A}^*: f \to \mathscr{A}^*(f), \forall f \in \mathfrak{L}^*$. 这时$\mathscr{A}^*$ 为 $\mathfrak{L}^*$ 上线性变换，称为$\mathscr{A}$的**对偶变换**或**共轭变换**. 在 $\mathfrak{L}$ 中任取一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，记$\mathscr{A}$对应的方阵为 $A=(a_{ij})$，再在 $\mathfrak{L}^*$ 中取$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$的对偶基 $f_1,f_2,\cdots,f_n$，则共轭变换 $\mathscr{A}^*$ 在对偶基 $f_1,f_2,\cdots,f_n$ 下对应的方阵为 A'.

证 由定义关系 $(\mathscr{A}^*(f))(\alpha)=f(\mathscr{A}(\alpha))$, $\forall \alpha \in \mathfrak{L}$, $f \in \mathfrak{L}^*$ 及直接计算可知 $\mathscr{A}^*$ 为 $\mathfrak{L}^*$ 上线性变换.

设 $\mathscr{A}(\alpha_i)=\sum_{j=1}^{n} a_{ji}\alpha_j$, $\mathscr{A}^*(f_i)=\sum_{j=1}^{n} b_{ji}f_j$, $i=1,2,\cdots,n$, 其中 $B=(b_{ij})$为$\mathscr{A}^*$ 在对偶基 $f_1,f_2,\cdots,f_n$ 下对应的方阵. 今

$$(\mathscr{A}^*(f_i))(\alpha_j)=f_i(\mathscr{A}(\alpha_j)), \qquad i,j=1,2,\cdots,n.$$

所以有

$$\sum_{k=1}^{n} b_{ki}f_k(\alpha_j)=f_i\left(\sum_{l=1}^{n} a_{lj}\alpha_l\right).$$

这证明了

$$\sum_{k=1}^{n} b_{kj}\delta_{jk}=\sum_{l=1}^{n} a_{lj}\delta_{li},$$

或 $b_{ji}=a_{ij}$, $i,j=1,2,\cdots,n$. 此即 $B=A'$. 证完.

定义 n 维线性空间 $\mathfrak{L}$ 上函数 $f(\alpha,\beta)$称为**双线性函数**，如果

有

$$f(a\alpha+\tilde{a}\tilde{\alpha},\beta)=af(\alpha,\beta)+\tilde{a}f(\tilde{\alpha},\beta),$$
$$f(\alpha,b\beta+\tilde{b}\tilde{\beta})=bf(\alpha,\beta)+\tilde{b}f(\alpha,\tilde{\beta}),$$

$\forall\alpha,\tilde{\alpha},\beta,\tilde{\beta}\in\mathfrak{L}$，$a,\tilde{a},b,\tilde{b}$ 为数.

定理 8.1.5 在 n 维线性空间 $\mathfrak{L}$ 中取定基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 设 f 为 $\mathfrak{L}$ 上双线性函数. 记

$$F=\begin{pmatrix} f(\alpha_1,\alpha_1) & \cdots & f(\alpha_1,\alpha_n) \\ \vdots & & \vdots \\ f(\alpha_n,\alpha_1) & \cdots & f(\alpha_n,\alpha_n) \end{pmatrix},$$

称为双线性函数 f 在基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 下的方阵表示. 则对 $\mathfrak{L}$ 中任两向量 $\alpha=\sum_{i=1}^{n}x_i\alpha_i$，$\beta=\sum_{j=1}^{n}y_j\alpha_j$，有

$$f(\alpha,\beta)=x'Fy$$

其中

$$x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\qquad y=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

反之，任取 n 阶方阵 F，则 $x'Fy=f\left(\sum_{i=1}^{n}x_i\alpha_i,\ \sum_{j=1}^{n}y_j\alpha_j\right)$ 定义了 $\mathfrak{L}$ 上一个双线性函数，所以双线性函数 f 与 n 阶方阵 F 之间的对应为一一对应.

证 今

$$f(\alpha,\beta)=f\left(\sum_{i=1}^{n}x_i\alpha_i,\ \sum_{j=1}^{n}y_j\alpha_j\right)=\sum_{i,j=1}^{n}x_iy_jf(\alpha_i,\alpha_j)=x'Fy.$$

反之，任取 n 阶方阵 $F=(f_{ij})$，由定义可知 $f\left(\sum_{i=1}^{n}x_i\alpha_i,\ \sum_{j=1}^{n}y_j\alpha_j\right)$ $=x'Fy$ 为 $\mathfrak{L}$ 上双线性函数. 为了证明 $f\to F$ 为一一对应，只要证 $\mathfrak{L}$ 上双线性函数 f，g 有 $f=g$ 当且仅当 $f(\alpha_i,\alpha_j)=g(\alpha_i,\alpha_j)$，

$i,j=1,2,\cdots,n$. 事实上,记F和G分别为,双线性函数f,g在基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下的方阵表示,由$f(\alpha,\beta)=x'Fy=x'Gy=g(\alpha,\beta)$,$\forall\alpha,\ \beta\in\Omega$ 可知当n阶方阵 $F=G$,则有双线性函数$f=g$. 反之显然. 证完.

定理 8.1.6 在n维线性空间Ω中取定基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 及$\{\beta_1,\beta_2,\cdots,\beta_n\}$,记基变换对应的$n$阶非异方阵为$P=(p_{ij})$. 设$f$为$\Omega$上双线性函数,它们在基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$及$\{\beta_1,\beta_2,\cdots,\beta_n\}$时对应的$n$阶方阵分别为$F,G$,则有

$$G=P'FP.$$

证 由G的定义可知,它的第i行,第j列元素为

$$\begin{aligned}f(\beta_i,\beta_j)&=f\left(\sum_{k=1}^{n}p_{ki}\alpha_k,\ \sum_{l=1}^{n}p_{lj}\alpha_l\right)\\&=\sum_{k,l=1}^{n}p_{ki}f(\alpha_k,\alpha_l)p_{lj},\end{aligned}$$

其中$f(\alpha_k,\alpha_l)$为n阶方阵F的第k行,第l列元素,将上式写成方阵乘积的形式,则有$G=P'FP$. 证完.

于是自然地引进

定义 n阶方阵A和B称为**相合的**,如果存在n阶非异方阵P,使得

$$B=P'AP.$$

定理 8.1.7 n阶方阵间的相合关系为等价关系。

证 在相合定义中取P为n阶单位方阵,所以证明了A和A相合,即证明了反身性. 设A和B相合,即存在n阶非异方阵P,使得$B=P'AP$. 于是$A=(P^{-1})'B(P^{-1})$,这证明了B和A相合,所以证明了对称性. 最后,若n阶方阵A,B,C有A和B相合,B和C相合,即存在n阶非异方阵P,Q,使得$B=P'AP$, $C=Q'BQ$. 于是$C=Q'(P'AP)Q=(PQ)'A(PQ)$. 这证明了A和C相合,所以

证明了传递性. 至此证明了相合关系为等价关系. 证完.

由相合定义可知

定理8.1.8 两个 n 阶方阵 A 和 B 相合当且仅当在 n 维线性空间 $\mathfrak{L}$ 上存在双线性函数 f, 使得 A 和 B 为 f 在不同基下对应的方阵，所以所有互相相合的方阵构成的集合(即等价类)和双线性函数间有一个自然的一一对应关系。

上面定理实际上给出了方阵相合这个概念的几何解释，因为双线性函数的定义与空间有关，与基底选取无关。

为了在第十章、第十一章和第十二章中有用，在下面引进一些重要的双线性函数.

定义 n 维线性空间 $\mathfrak{L}$ 上双线性函数 f 称为**对称**(**斜对称**)的，如果有 $f(\alpha,\beta)=f(\beta,\alpha)\,(f(\alpha,\beta)=-f(\beta,\alpha))$, $\forall\alpha,\beta\in\mathfrak{L}$.

定理 8.1.9 n 维线性空间 $\mathfrak{L}$ 上双线性函数为对称 (斜对称的当且仅当在 $\mathfrak{L}$ 中存在一组基，使得在这组基下它对应的方阵表) 示为对称(斜对称)方阵.

证 在 $\mathfrak{L}$ 中任取一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 设 f 为对称双线性函数，则有 $f(\alpha_i,\alpha_j)=f(\alpha_j,\alpha_i)$. 于是

$$F=\begin{pmatrix} f(\alpha_1,\alpha_1) & \cdots & f(\alpha_1,\alpha_n) \\ \vdots & & \vdots \\ f(\alpha_n,\alpha_1) & \cdots & f(\alpha_n,\alpha_n) \end{pmatrix}$$

为对称方阵. 反之，若 F 对称，即 $f(\alpha_i,\alpha_j)=f(\alpha_j,\alpha_i)$, $1\leqslant i,j\leqslant n$, 自然 $f(\alpha,\beta)=f(\beta,\alpha)$, 即 f 为对称双线性函数. 在斜对称情形证明是相同的. 证完.

下面定理是比较重要的.

引理 8.1.1 设 A 为 n 阶方阵，则

$$S=\frac{1}{2}(A+A'),\qquad K=\frac{1}{2}(A-A')$$

分别为对称方阵及斜对称方阵，且有

$$A=S+K.$$

进而，若存在对称方阵 S_1，斜对称方阵 K_1，使得

$$A=S_1+K_1,$$

则有 $S_1=S$, $K_1=K$，即 n 阶方阵 A 能唯一地分解为一个对称方阵 S（称为 A 的**对称部分**）和一个斜对称方阵 K（称为 A 的**斜对称部分**）之和

证 显然，$S'=S, K'=-K$，下面证唯一性，今若 $A=S+K=S_1+K_1$，于是 $S-S_1=K_1-K$。因此

$$\begin{aligned} S-S_1&=S'-S_1'=(K_1-K)' \\ &=K_1'-K'=-K_1+K=S_1-S \end{aligned}$$

这证明了 $S=S_1$，所以 $K=K_1$. 证完.

定理 8.1.10 n 维线性空间 Ω 上双线性函数 f 可唯一地分解为一个对称双线性函数 s 和一个斜对称双线性函数 k 之和，且在 Ω 中取定基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 后，则 f, s, k 分别对应方阵为

$$A, S=\frac{1}{2}(A+A'), K=\frac{1}{2}(A-A').$$

证 记 $s(\alpha,\beta)=\frac{1}{2}f(\alpha, \beta)+\frac{1}{2}f(\beta,\alpha), k(\alpha, \beta)$

$$=\frac{1}{2}f(\alpha, \beta)-\frac{1}{2}f(\beta,\alpha).$$

不难证明 s 为对称双线性函数，k 为斜对称双线性函数. 显然 $f=s+k$. 在 Ω 中取定基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，设 f 对应 n 阶方阵为 A，则 s 和 k 分别对应 $\frac{1}{2}(A+A')=S$ 及 $\frac{1}{2}(A-A')=K$. 由于若 $f=s_1+k_1$，其中 s_1 对称，k_1 斜对称，它们分别对应 n 阶方阵 A,S_1,K_1，显然 S_1,K_1 分别为对称及斜对称方阵，又 $A=S_1+K_1$. 由引理 8.1.1，所以 $S_1=\frac{1}{2}(A+A')$，$K_1=\frac{1}{2}(A-A')$. 这证明了 $s_1=s, k_1=k$，即

分解唯一．证完．

定义 设 f 为 n 维线性空间 Ω 上双线性函数，则

$$\varphi(\alpha)=f(\alpha,\alpha),\quad \forall\alpha\in\Omega$$

为 Ω 上的函数，称为**二次型**．当 Ω 为实线性空间时，实二次型 $\varphi(\alpha)$ 称为**定正**（**半定正**、**定负**、**半定负**）的，如果 $\varphi(\alpha)>0(\geqslant 0,<0,\leqslant 0)$，$\forall\alpha\neq 0,\alpha\in\Omega$．

定理 8.1.11 在 n 维线性空间 Ω 中取定一组基 $\{\alpha_1,\ \alpha_2,\cdots,\alpha_n\}$，则二次型 $\varphi(\alpha)$ 有坐标表达式

$$\varphi(\alpha)=x'Sx,$$

其中 S 为对称方阵，且上述表达式唯一．

证 今 $f(\alpha,\beta)=x'Fy$，于是

$$\begin{aligned}\varphi(\alpha)=f(\alpha,\alpha)=x'Fx&=\frac{1}{2}(x'Fx+x'F'x)\\&=x'\left(\frac{1}{2}(F+F')\right)x=x'Sx,\end{aligned}$$

其中 S 为 F 的对称部分，再若 $x'Sx=x'S_1x$，其中 S，S_1 都是 n 阶对称方阵，于是 $x'(S-S_1)x=0$，其中 $S-S_1$ 也对称，比较 $x_1,x_2,\cdots,x_n$ 的各项系数，便证明了 $S-S_1=0$．证完．

定义 n 阶实对称方阵 S 称为**定正**（**半定正**、**定负**、**半定负**）的，记作 $S>0(\geqslant 0,<0,\leqslant 0)$，如果二次型 $x'Sx>0(\geqslant 0,<0,\ \leqslant 0)$ 对一切 $0\neq x\in V_n$ 成立、其中 V_n 为所有 $n\times 1$ 实矩阵构成的 n 维实线性空间．

注意，由于当 S 为 n 阶复方阵，x 为 $n\times 1$ 复方阵，则 $x'Sx$ 仍为复数，所以上述定义失效．在复的情形，我们从下面角度推广．

引理 8.1.2 设 A 为 n 阶复方阵，则 $\frac{1}{2}(A+\bar{A}')=H$，$\frac{1}{2}(A-\bar{A}')=K$ 分别为 Hermite 方阵及斜 Hermite 方阵，且 $A=H+K$，即 A 能分解为 Hermite 方阵 H（称为 Hermite **部分**）

和斜 Hermite 方阵 K(称为斜 Hermite 部分)之和. 且上述分解唯一.

证 分解的存在性显然，下面证唯一性. 设 A 为 n 阶方阵，且 $A=H_1+K_1$，其中 H_1 为 Hermite 方阵，K_1 为斜 Hermite 方阵. 于是 $\bar{A}'=\bar{H}_1'+\bar{K}_1'=H_1-K_1$. 因此 $A+\bar{A}'=2H_1, A-\bar{A}'=2K_1$，这证明了 $H_1=H, K_1=K$，所以证明了分解的唯一性. 证完.

定义 n 维复线性空间 Ω 上复值函数 $h(\alpha,\beta)$ 称为 Ω 上 Hermite **双线性函数**，如果它适合条件:

(1) $h(a\alpha+\tilde{a}\tilde{\alpha},\beta)=ah(\alpha,\beta)+\tilde{a}h(\tilde{\alpha},\beta)$,

(2) $h(\alpha,b\beta+cr)=\bar{b}h(\alpha,\beta)+\bar{c}h(\alpha,r)$,

其中 $a,\tilde{a},b,c$ 为复数，$\alpha,\tilde{\alpha},\beta,r\in\Omega$.

定理 8.1.12 在 n 维复线性空间 Ω 中任取基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$，给定Hermite 双线性函数 $h(\alpha,\beta)$，则

$$h(\alpha,\beta)=x'A\bar{y},$$

其中 $\alpha=\sum_{j=1}^{n}x_j\alpha_j,\ \beta=\sum_{k=1}^{n}y_k\alpha_k$，又

$$x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\quad y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix},\quad A=\begin{pmatrix}h(\alpha_1,\alpha_1)&\cdots&h(\alpha_1,\alpha_n)\\\vdots&&\vdots\\h(\alpha_n,\alpha_1)&\cdots&h(\alpha_n,\alpha_n)\end{pmatrix}.$$

反之，任给 n 阶复方阵 A，则

$$h\left(\sum_{j=1}^{n}x_j\alpha_j,\sum_{k=1}^{n}y_k\alpha_k\right)=x'A\bar{y}$$

为 Ω 上 Hermite 双线性函数，所以 Hermite 双线性函数和 n 阶复方阵间有如下的一一对应: $h\to A$.

证 今

$$h(\alpha,\ \beta)=h\left(\sum_{j=1}^{n}x_j\alpha_j,\ \sum_{k=1}^{n}y_k\alpha_k\right)=\sum_{j,\,k=1}^{n}x_jh(\alpha_j,\alpha_k)\bar{y}_k=x'A\bar{y}.$$

反之,由定义可以验证$h\left(\sum_{j=1}^{n}x_j\alpha_j,\ \sum_{k=1}^{n}y_k\alpha_k\right)=x'A\bar{y}$ 为 Hermite 双线性函数. 最后,为了证 $h\to A$ 为一一对应,只要证 $x'A\bar{y}=x'B\bar{y}$, $\forall x,y\in V_n(C)$,则 $A=B$,其中 $V_n(C)$为所有 $n\times 1$ 复矩阵构成的 n 维复线性空间. 比较同类项系数可知断言成立. 证完.

定理8.1.13 在 n 维复线性空间 $\mathfrak{L}$ 中任取两组基$\{\alpha_1,\ \alpha_2,\cdots,\ \alpha_n\}$和$\{\beta_1,\beta_2,\cdots,\beta_n\}$. 记$\beta_i=\sum_{j=1}^{n}p_{ji}\alpha_j$, $1\leqslant i\leqslant n$,其中$P=(p_{ij})$为 n 阶非异复方阵, 记 h 为 $\mathfrak{L}$ 上 Hermite 双线性函数, 它在基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$和$\{\beta_1,\beta_2,\cdots,\beta_n\}$下对应的方阵分别为 n 阶复方阵 A 和 B. 则有

$$B=P'A\bar{P}$$

证 今

$$h(\beta_i,\beta_j)=h\left(\sum_{k=1}^{n}p_{ki}\alpha_k,\sum_{l=1}^{n}p_{lj}\alpha_l\right)=\sum_{k,\,l=1}^{n}p_{ki}h(\alpha_k,\alpha_l)\bar{p}_{lj}$$

写成方阵形式,即有 $B=P'A\bar{P}$. 证完.

定义 n 阶复方阵 A 和 B 称为**复相合的**, 如果存在 n 阶非异复方阵 P,使得 $B=P'A\bar{P}$.

由此定义,上面定理可以改写为

定理 8.1.14 n 维复线性空间上的 Hermite 双线性函数在不同基下对应的方阵是复相合的.

定理 8.1.15 复方阵的复相合关系是等价关系.

证 取 P 为 n 阶单位方阵,则有 $A=P'A\bar{P}$,即 A 和 A 复相合,这证明了复相合的反身性. 设 P 为 n 阶非异复方阵,且对 n 阶复方阵 A,B,有 $B=P'A\bar{P}$,于是有 $A=(P^{-1})'B(\overline{P^{-1}})$,即若 A 和 B 复相

合，则 B 和 A 复相合，这证明了对称性. 最后，若 n 阶复方阵 A, B, C 有 A 和 B 复相合，B 和 C 复相合，即若存在 n 阶非异复方阵 P, Q，使得 $B=P'A\overline{P}, C=Q'B\overline{Q}$，于是有 $C=Q'B\overline{Q}=Q'P'A\overline{P}\overline{Q}=(PQ)'A(\overline{PQ})$，这证明了 A 和 C 复相合，所以证明了传递性，因此证明了复相合为等价关系. 证完.

下面考虑两类特殊的 Hermite 双线性函数.

定义 n 维复线性空间 Ω 上Hermite 双线性函数 $h(\alpha,\beta)$ 称为 Hermite 的（斜 Hermite 的），如果有 $h(\alpha,\beta)=\overline{h(\beta,\alpha)}$ $(h(\alpha,\beta)=-\overline{h(\beta_1,\alpha)})$，$\forall\alpha,\beta\in\Omega$.

于是显然有

定理 8.1.16 在 n 维复线性空间 Ω 中任取一组基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$，则Hermite的Hermite双线性函数对应的方阵为 n 阶Hermite方阵，斜 Hermite 的 Hermite 双线性函数对应的方阵为 n 阶斜 Hermite 方阵.

且由引理 8.1.2 有

定理 8.1.17 n 维复线性空间 Ω 上Hermite 双线性函数能唯一地表成一个 Hermite 的 Hermite 双线性函数和一个斜Hermite 的 Hermite 双线性函数之和. 换句话说，任一 n 阶复方阵能唯一地表成一个 Hermite 方阵和一个斜 Hermite 方阵之和.

定义 设 $h(\alpha,\beta)$ 为 n 维复线性空间 Ω 上Hermite的Hermite 双线性函数，则

$$\varphi(\alpha)=h(\alpha,\alpha),\quad \forall\alpha\in\Omega$$

为 Ω 上的函数，称为 Ω 上 Hermite 型、它是 Ω 上实值函数. 设 $\varphi(\alpha)>0(\geqslant 0, <0, \leqslant 0)$，$\forall\alpha\in\Omega$，$\alpha\neq 0$，则 φ 称为 Ω 上定正（半定正、定负、半定负）Hermite 型.

定理 8.1.18 在 n 维复线性空间 Ω 中取定基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，则 Hermite 型 $\varphi(\alpha)$ 有坐标表达式

$$\varphi(\alpha)=x'H\bar{x},$$

其中$\alpha=\sum_{j=1}^{n}x_j\alpha_j$, 又 H 为 Hermite 方阵. 且上述表达式唯一, 即 Hermite 型和 Hermite 方阵间是一一对应的.

证 今 $h(\alpha,\beta)$为 n 维复线性空间 $\mathfrak{L}$ 上 Hermite 的 Hermite 双线性函数,则 $\varphi(\alpha)=h(\alpha,\alpha)$为 $\mathfrak{L}$ 上 Hermite 型. 在 $\mathfrak{L}$ 中取定基,则有

$$\varphi(\alpha)=h(\alpha,\alpha)=x'A\bar{x}.$$

由$\overline{h(\alpha,\alpha)}=h(\alpha,\alpha)$,所以$\overline{(x'A\bar{x})'}=x'\bar{A}'\bar{x}$. 于是

$$\varphi(\alpha)=h(\alpha,\alpha)=x'H\bar{x},\quad H=\frac{1}{2}(A+\bar{A}').$$

再证唯一性. 今若 H_1, H 为 n 阶 Hermite 方阵, 且有 $x'H\bar{x}=x'H_1\bar{x}$, $\forall x\in V_n(C)$, 其中 $V_n(C)$为所有 $n\times1$ 复矩阵构成的线性空间. 于是 $x'(H-H_1)\bar{x}=0$, $\forall x\in V_n(C)$. 比较 $x_1,\cdots,x_n$ 的各种二次项 $x_i\bar{x}_j$ 之系数,便证明了 $H_1=H$. 证完.

定义 n 阶 Hermite 方阵 H 称为**定正**(**半定正**, **定负**, **半定负**)的,如果它唯一决定的Hermite 型$x'H\bar{x}$ 是定正(半定正,定负,半定负)的. 即对一切 $0\neq x\in V_n(C)$,则有 $x'H\bar{x}>0(\geqslant0,<0,\leqslant0)$. 这时记作 $H>0(\geqslant0,<0,\leqslant0)$.

习题 8.1

1. 试证:实线性空间上的有限个线性函数的平方和为半定正二次型

2. 设 $f(\alpha,\beta)$为实线性空间 $\mathfrak{L}$ 上双线性函数,记

$$\ker(f)=\{\alpha\in\mathfrak{L}\,|\,f(\alpha,\beta)=0,\ \forall\beta\in\mathfrak{L}\}.$$

试证: $\ker(f)$为 $\mathfrak{L}$ 的子空间,称为 f 的核. 当 $\ker(f)=0$,则称 $\mathfrak{L}$ 上双线性函数 f 为非退化的. 试证: f 为非退化的当且仅当它对应的方阵是非异的. 在 $\mathfrak{L}$ 中任取向量 α,β,如果 $f(\alpha,\beta)=0$, 则称 α, β 关于双线性函数 f 是正交的. 试证: 在 $\mathfrak{L}$ 中任给子空间 $\mathfrak{L}_1$,则

$$\mathfrak{L}_1^{\perp}=\{\alpha\in\mathfrak{L}\mid f(\alpha,\beta)=0,\ \forall\beta\in\mathfrak{L}_1\}$$

为 $\mathfrak{L}$ 的子空间，称为 $\mathfrak{L}_1$ 的关于 f 的正交补.

设 $h(\alpha,\beta)$ 为复线性空间 $\mathfrak{L}$ 上 Hermite 双线性函数，和上面 f 一样可定义核、非退化、正交补. 试证：上述性质也都成立.

3. 记 $f(\alpha,\beta)$ 为实线性空间 $\mathfrak{L}$ 上对称双线性函数，记

$$\mathfrak{N}_0=\{\alpha\in\mathfrak{L}\mid f(\alpha,\alpha)=0\}$$

则 $\mathfrak{N}_0$ 中子空间称为**全迷向子空间**.

(i) 试证：$\mathfrak{L}$ 中子空间 $\mathfrak{M}$ 为全迷向子空间当且仅当 $\mathfrak{M}\subset\mathfrak{M}^{\perp}$,

(ii) 设 $\mathfrak{M}_1,\mathfrak{M}_2$ 为 $\mathfrak{L}$ 中两全迷向子空间，$\mathfrak{P}_1$, $\mathfrak{P}_2$ 为子空间. 设 $\mathfrak{M}_i$ 有子空间直接和分解

$$\mathfrak{M}_i=\mathfrak{P}_i\oplus(\mathfrak{M}_1\cap\mathfrak{M}_2),\ i=1,2.$$

试证：$\mathfrak{M}_1\cap\mathfrak{M}_2^{\perp}$ 有子空间直接和分解

$$\mathfrak{M}_1\cap\mathfrak{M}_2^{\perp}=(\mathfrak{M}_1\cap\mathfrak{M}_2)\oplus(\mathfrak{P}_1\cap\mathfrak{P}_2^{\perp}).$$

§8.2 多重线性函数和张量

定义 设 $\mathfrak{L}$ 为 n 维线性空间，给定 s 个互相独立的变向量 $\xi_1,\xi_2,\cdots,\xi_s$, t 个互相独立的变线性函数 $g_1,g_2,\cdots,g_t$. 设这 $s+t$ 个变元的函数

$$f(\xi_1,\cdots,\xi_s;\ g_1,\cdots,g_t)$$

适合条件：在其中任意固定 $s+t-1$ 个变元，则为余下变元的线性函数. 这时 f 称为**多重线性函数**，f 又称为 $(s,\ t)$ **型张量**，即 s 阶共变，t 阶反变张量. 它们全体构成集合 $\mathscr{D}_t^s$. 特别，$\mathscr{D}_0^0$ 约定为所有纯量构成的集合.

显然，在 $\mathscr{D}_t^s$ 中可引进加法及纯量乘积如下：

(1) 加法，$\forall f,g\in\mathscr{D}_t^s$,

$$(f+g)(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)=f(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)+g(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t),$$

(2) 纯量乘积，$\forall f\in\mathscr{D}_t^s$, a 为纯量

$$(af)(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)=af(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t).$$

由定义不难验证 $\mathscr{D}_t^s$ 为线性空间.

引理 8.2.1 $\mathscr{D}_t^s$ 为 n^{s+t} 维线性空间.

证 在 Ω 中取定一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 记 $f_1,f_2,\cdots,f_n$ 为 Ω 的对偶空间 Ω^* 中的一组对偶基. 于是有

$$f_i(\alpha_j)=\delta_{ij},\quad i,j=1,2,\cdots,n.$$

记 s 个变向量

$$\xi_i=\sum_{j=1}^{n}x_{ij}\alpha_j,\quad i=1,2,\cdots,s$$

记 t 个变线性函数

$$g_k=\sum_{l=1}^{n}y_{kl}f_l,\quad k=1,2,\cdots,t.$$

任取 $f\in\mathscr{D}_t^s$, 于是

$$f(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)=\sum_{i_1,\cdots,i_s=1}^{n}\ \sum_{j_1,\cdots,j_t=1}^{n}x_{1i_1}x_{2i_2}\cdots x_{si_s}$$

$$\times y_{1j_1}y_{2j_2}\cdots y_{tj_t}f(\alpha_{i_1},\cdots,\alpha_{i_s};f_{j_1},\cdots,f_{j_t}).$$

记常量

$$f(\alpha_{i_1},\cdots,\alpha_{i_s};f_{j_1},\cdots,f_{j_t})=a_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t},1\leqslant i_1,\cdots,i_s,j_1,\cdots,j_t\leqslant n$$

所以(s,t)型张量可以看作 n^{s+t} 个独立自变量 x_{ij}, $1\leqslant i\leqslant s$, $1\leqslant j\leqslant n$ 及 $y_{ij},1\leqslant i\leqslant t,1\leqslant j\leqslant n$ 的函数. 我们称 n^{s+t} 个数

$$a_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t},1\leqslant i_1,\cdots,i_s,j_1,\cdots,j_t\leqslant n$$

为(s, t)型张量 f 在基 α_1, $\cdots$, α_n 下的坐标. 显然, 它唯一决定 f. 反之, 任给 n^{s+t} 个数 $a_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t},1\leqslant i_1,\cdots,i_s,j_1,\cdots,j_t\leqslant n$, 考虑

$$\sum_{i_1,\cdots,i_s=1}^{n}\ \sum_{j_1,\cdots,j_t=1}^{n}\ a_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t}\ x_{1i_1}x_{2i_2}\cdots x_{si_s}y_{1j_1}y_{2j_2}\cdots y_{tj_t}$$

在 Ω 中取定基 $\alpha_1,\cdots,\alpha_n$, 则它定义了 Ω 上(s,t)型张量.

由此可见, $\mathscr{D}_t^s$ 中元素由且仅由 n^{s+t} 个数唯一确定. 因此我们取定指标 $i_1,\cdots,i_s,j_1,\cdots,j_t$, 使得它们取 $1,2,\cdots,n$之一. 取一组数

$$a_{k_1 k_2\cdots k_s}^{l_1 l_2\cdots l_t}=\delta_{i_1k_1}\delta_{i_2k_2}\cdots\delta_{i_sk_s}\delta_{j_1l_1}\delta_{j_2l_2}\cdots\delta_{j_tl_t}$$

于是定义了 $\mathscr{D}_t^s$ 中元素 $f_{i_1 i_2\cdots i_s}^{j_1 j_2\cdots j_t}$, $1\leqslant i_1,\cdots,i_s,j_1,\cdots,j_t\leqslant n$. 即有

$$f_{i_1 i_2\cdots i_s}^{j_1 j_2\cdots j_t}(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)=x_{1i_1}x_{2i_2}\cdots x_{si_s}y_{1j_1}y_{2j_2}\cdots y_{tj_t}.$$

我们来证 $\{f_{i_1 i_2\cdots i_s}^{j_1 j_2\cdots j_t},1\leqslant i_1,\cdots,i_s,j_1,\cdots,j_t\leqslant n\}$ 为 $\mathscr{D}_t^s$ 中一组基. 事实上，任取 $f\in\mathscr{D}_t^s$，则

$$f(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)=\sum f(\alpha_{i_1},\cdots,\alpha_{i_s};f_{j_1},\cdots,f_{j_t})$$
$$\times f_{i_1 i_2\cdots i_s}^{j_1 j_2\cdots j_t}(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t).$$

这证明了 $f_{i_1 i_2\cdots i_s}^{j_1 j_2\cdots j_t}$, $1\leqslant i_1,\cdots,i_s,j_1,\cdots,j_t\leqslant n$ 线性生成 $\mathscr{D}_t^s$. 最后证它们线性无关. 事实上，若

$$\sum_{i_1,\cdots,i_s=1}^{n}\sum_{j_1,\cdots,j_t=1}^{n}a_{i_1 i_2\cdots i_s}^{j_1 j_2\cdots j_t}f_{i_1 i_2\cdots i_s}^{j_1 j_2\cdots j_t}=0,$$

此即

$$\sum_{i_1,\cdots,i_s=1}^{n}\sum_{j_1,\cdots,j_t=1}^{n}a_{i_1 i_2\cdots i_s}^{j_1 j_2\cdots j_t}x_{1i_1}x_{2i_2}\cdots x_{si_s}y_{1j_1}y_{2j_2}\cdots y_{tj_t}=0.$$

这证明了 $a_{i_1 i_2\cdots i_s}^{j_1 j_2\cdots j_t}=0$, $\forall\, 1\leqslant i_1,\cdots,i_s,j_1,\cdots,j_t\leqslant n$. 所以断言成立. 因此 $\mathscr{D}_t^s$ 的维数为 n^{s+t}. 证完.

作形式直接和

$$\mathscr{D}=\bigoplus_{s,t=0}^{\infty}\mathscr{D}_t^s,$$

使得 $\mathscr{D}$ 中任一元素可表为

$$F=\sum_{s,t=0}^{\infty}f_t^s,$$

其中 $f_t^s\in\mathscr{D}_t^s$，且上述表法唯一，即若

$$\sum_{s,t=0}^{\infty}f_t^s=\sum_{s,t=0}^{\infty}g_t^s,\quad \forall f_t^s,g_t^s\in\mathscr{D}_t^s,$$

则必须 $f_t^s=g_t^s$, $0\leqslant s,\ t<\infty$. 于是 $\mathscr{D}$ 构成无限维线性空间，其中

加法及纯量乘积分别定义为

$$\sum_{s,t=0}^{\infty} f_t^s + \sum_{s,t=0}^{\infty} g_t^s = \sum_{s,t=0}^{\infty} (f_t^s + g_t^s),$$

$$a\sum_{s,t=0}^{\infty} f_t^s = \sum_{s,t=0}^{\infty} (af_t^s).$$

重要的是引进下面张量积的概念.

(1) 任取 $f\in\mathscr{D}_t^s, g\in\mathscr{D}_q^p$, 定义 f 和 g 的**张量积** $f\otimes g$ 如下;首先, $f\otimes g\in\mathscr{D}_{t+q}^{s+p}$; 其次,

$$(f\otimes g)(\xi_1,\cdots,\xi_s,\xi_{s+1},\cdots,\xi_{s+p};g_1,\cdots,g_t,g_{t+1},\cdots,g_{t+q})$$
$$=f(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)g(\xi_{s+1},\cdots,\xi_{s+p};g_{t+1},\cdots,g_{t+q}).$$

由定义不难验证上述张量积定义有结合律,无交换律,有单位元素 $1\in\mathscr{D}_0^0$.

(2) 任取 $\mathscr{D}$ 中两元素 $F=\sum_{s,t=0}^{\infty} f_t^s$, $G=\sum_{s,t=0}^{\infty} g_t^s$, 定义 F 和 G 的张量积 $F\otimes G$ 如下: 首先, $F\otimes G\in\mathscr{D}\otimes\mathscr{D}=\mathscr{D}$; 其次,

$$F\otimes G=\left(\sum_{s,t=0}^{\infty} f_t^s\right)\otimes\left(\sum_{p,q=0}^{\infty} g_q^p\right)=\sum_{s,t,p,q=0}^{\infty}(f_t^s\otimes g_q^p),$$

所以

$$F\otimes G=\sum_{u,v=0}^{\infty}\left(\sum_{s+p=u}\sum_{t+q=v} f_t^s\otimes g_q^p\right)=\sum_{u,v=0}^{\infty} h_v^u,$$

其中

$$h_v^u=\sum_{s+p=u}\sum_{t+q=v} f_t^s\otimes g_q^p,\ \ 0\leqslant u,v<\infty.$$

由定义,显然 $\mathscr{D}$ 中张量积有结合律,无交换律,且有单位元素 $1\in\mathscr{D}_0^0$.

定义 线性空间 $\mathfrak{L}$ 中如果可以引进乘法,它适合

(1) 乘法结合律

$$\alpha(\beta\gamma)=(\alpha\beta)\gamma, \quad \forall\alpha,\beta,\gamma\in\mathfrak{L};$$

（2）加乘分配律

$$\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma, \quad (\beta+\gamma)\alpha=\beta\alpha+\gamma\alpha, \quad \forall\alpha,\beta,\gamma\in\mathfrak{L};$$

（3）乘法和纯量乘积间有关系

$a(\alpha\beta)=(a\alpha)\beta=\alpha(a\beta)$，$\forall\alpha,\beta\in\mathfrak{L}$，$a$ 为纯量，则 $\mathfrak{L}$ 称为**结合代数**．$\mathfrak{L}$ 中子空间 $\mathfrak{L}_1$ 称为**子代数**，如果任取 $\alpha,\beta\in\mathfrak{L}_1$，则 $\alpha\beta\in\mathfrak{L}_1$．显然它仍为结合代数．

结合代数是代数学中最重要的"代数"，例如：所有 n 阶方阵，在加法、乘法、纯量乘积下构成 n^2 维结合代数．关于结合代数的研究属于抽象代数这一学科．我们引进这个概念，是为了给出

定理 8.2.1 无限维线性空间 $\mathscr{D}$ 在张量积下构成结合代数，称为线性空间 $\mathfrak{L}$ 上**张量代数**．

由 $\mathscr{D}$ 之张量积定义直接验证便可证明定理，在此略去．

张量代数是近代微分几何的重要工具，它首先是由 Einstein 在研究相对论时应用到物理学上去的，从而推动了近代微分几何的发展．但是 Einstein 是用张量的坐标来表示张量，而坐标和线性空间 $\mathfrak{L}$ 上的基底有关，为此需要给出在不同基底下坐标间的关系．

引理 8.2.2 设 $\mathfrak{L}$ 为 n 维线性空间，$\mathfrak{L}^*$ 为其对偶空间．$(\mathfrak{L}^*)^*$ 为 $\mathfrak{L}^*$ 的对偶空间．任取 $\Phi\in(\mathfrak{L}^*)^*$，则唯一存在 $\mathfrak{L}$ 中向量 α，使得

$$\Phi(\varphi)=\varphi(\alpha), \quad \forall\varphi\in\mathfrak{L}^*.$$

于是建立了 $(\mathfrak{L}^*)^*$ 到 $\mathfrak{L}$ 上的线性同构 $\Phi\to\alpha$．在这意义下，我们改记 Φ 为 α，即视 $(\mathfrak{L}^*)^*$ 为 $\mathfrak{L}$，于是有

$$\alpha(\varphi)=\varphi(\alpha), \quad \forall\varphi\in\mathfrak{L}^*, \quad \alpha\in\mathfrak{L}.$$

证 在 $\mathfrak{L}$ 中取其 $\{\alpha_1.\cdots,\alpha_n\}$，在 $\mathfrak{L}^*$ 中取对偶基 $\{f_1,\cdots,f_n\}$．在 $(\mathfrak{L}^*)^*$ 中取关于 $\{f_1,\cdots,f_n\}$ 之对偶基 $\{\Phi_1,\cdots,\Phi_n\}$．于是有

$$f_i(\alpha_j)=\delta_{ij}, \quad \Phi_i(f_j)=\delta_{ij}, \quad 1\leqslant i,j\leqslant n.$$

任取 $\Phi\in(\mathfrak{L}^*)^*$，则有

$$\Phi=\sum_{j=1}^{n}a_j\Phi_j.$$

在 $\mathfrak{L}$ 中取向量

$$\alpha=\sum_{j=1}^{n}a_j\alpha_j.$$

显然 $\Phi\to\alpha$ 为线性空间 $(\mathfrak{L}^*)^*$ 到 $\mathfrak{L}$ 上的线性映射. 今任取 $\varphi\in\mathfrak{L}^*$，记 $\varphi=\sum_{j=1}^{n}x_jf_j$，则

$$\begin{aligned}\Phi(\varphi)&=\sum_{j=1}^{n}a_j\Phi_j(\varphi)=\sum_{j=1}^{n}\sum_{k=1}^{n}a_jx_k\Phi_j(f_k)\\&=\sum_{j,k=1}^{n}a_jx_k\delta_{jk}=\sum_{j=1}^{n}a_jx_j,\\ \varphi(\alpha)&=\sum_{j=1}^{n}x_jf_j(\alpha)=\sum_{j=1}^{n}\sum_{k=1}^{n}x_ja_kf_j(\alpha_k)\\&=\sum_{j,k=1}^{n}x_ja_k\delta_{jk}=\sum_{j=1}^{n}a_jx_j.\end{aligned}$$

这证明了 $\Phi(\varphi)=\varphi(\alpha)$，$\forall\varphi\in\mathfrak{L}^*$. 余下要证：$\Phi\to\alpha$ 为线性同构. 线性性显然，所以只要证它一一到上. 今若 $\beta\in\mathfrak{L}$，且 $\Phi(\varphi)=\varphi(\beta)$，$\forall\varphi\in\mathfrak{L}^*$，则 $\beta=\alpha$. 事实上，记 $\beta=\sum_{j=1}^{n}b_j\alpha_j$，则同上计算

$$\sum_{j=1}^{n}a_jx_j=\Phi(\varphi)=\varphi(\beta)=\sum_{j=1}^{n}b_jx_j,$$

其中 $x_1,\cdots,x_n$ 可取任意值，所以证明了 $b_j=a_j$，$1\leqslant j\leqslant n$，即有 $\beta=\alpha$. 这证明了映射 $\Phi\to\alpha$ 为线性同构. 证完.

定理 8.2.2 在 n 维线性空间 $\mathfrak{L}$ 中取基 $\{\alpha_1,\cdots,\alpha_n\}$，在对偶空间 $\mathfrak{L}^*$ 中取对偶基 $\{f_1,\cdots,f_n\}$，则 $\mathfrak{D}_s^t$ 有基

$$f_{i_1}\otimes\cdots\otimes f_{i_s}\otimes\alpha_{j_1}\otimes\cdots\otimes\alpha_{j_t},\qquad 1\leqslant i_1,\cdots,i_s,j_1,\cdots,j_t\leqslant n.$$

于是

$$\mathscr{D}_t^s=\overbrace{\mathscr{D}_0^1\otimes\cdots\otimes\mathscr{D}_0^1}^{s}\otimes\overbrace{\mathscr{D}_1^0\otimes\cdots\otimes\mathscr{D}_1^0}^{t}$$

由 $\mathscr{D}_0^1=\mathfrak{L}^*,\mathscr{D}_1^0=\mathfrak{L}$，即有

$$\mathscr{D}_t^s=(\mathfrak{L}^*)^s\otimes(\mathfrak{L})^t$$

其中方幂按张量积来理解. 又线性空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 之张量积 $\mathfrak{L}_1\otimes\mathfrak{L}_2$ 理解为由 $\{\alpha_1\otimes\alpha_2\mid\forall\alpha_1\in\mathfrak{L}_1,\alpha_2\in\mathfrak{L}_2\}$ 线性生成的线性空间.

证 由定义可知 $\mathscr{D}_0^1=\mathfrak{L}^*$，$\mathscr{D}_1^0=(\mathfrak{L}^*)^*$. 由引理 8.2.2，所以 $\mathscr{D}_1^0=\mathfrak{L}$. 因此只要证明 $\mathscr{D}_t^s$ 中基

$$f_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t}=f_{i_1}\otimes\cdots\otimes f_{i_s}\otimes\alpha_{j_1}\otimes\cdots\otimes\alpha_{j_t},\qquad 1\leqslant i_1,\cdots,\ i_s,\ j_1,\cdots,j_t\leqslant n$$

就行了. 这里符号和引理 8.2.1 的证明中出现的符号相同.

已知 $\xi_i=\sum\limits_{j=1}^{n}x_{ij}\alpha_j$，$1\leqslant i\leqslant s$；$g_k=\sum\limits_{k=1}^{n}y_{kl}f_l$，$1\leqslant k\leqslant t$，则

$$f_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t}(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)=x_{1i_1}\cdots x_{si_s}y_{1j_1}\cdots y_{tj_t}.$$

而由张量积的定义可知

$$\begin{aligned}&(f_{i_1}\otimes\cdots\otimes f_{i_s}\otimes\alpha_{j_1}\otimes\cdots\otimes\alpha_{j_t})(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)\\&=f_{i_1}(\xi_1)\cdots f_{i_s}(\xi_s)\alpha_{j_1}(g_1)\cdots\alpha_{j_t}(g_t)\\&=f_{i_1}(\xi_1)\cdots f_{i_s}(\xi_s)g_1(\alpha_{j_1})\cdots g_t(\alpha_{j_t})\\&=x_{1i_1}\cdots x_{si_s}y_{1j_1}\cdots y_{tj_t}.\end{aligned}$$

所以证明了 $f_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t}=f_{i_1}\otimes\cdots\otimes f_{i_s}\otimes g_{j_1}\otimes\cdots\otimes g_{j_t}$. 证完.

定理 8.2.3 在 n 维线性空间 $\mathfrak{L}$ 中任取两组基 $\{\alpha_1,\cdots,\alpha_n\}$ 及 $\{\beta_1,\cdots,\beta_n\}$，基变换公式为

$$\beta_i=\sum_{j=1}^{n}p_{ji}\alpha_j,\quad 1\leqslant i\leqslant n,$$

其中 $P=(p_{ij})$ 为 n 阶非异方阵. 任取 $f\in\mathscr{D}_t^s$，设 f 在基 $\{\alpha_1,\cdots,\alpha_n\}$ 及 $\{\beta_1,\cdots\beta_n\}$ 下对应的坐标分别为

$$\{a_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t},\ 1\leqslant i_1,\cdots,i_s,j_1,\cdots,j_t\leqslant n\},$$

$$\{b_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t},\ 1\leqslant i_1,\cdots,i_s,j_1,\cdots,j_t\leqslant n\}.$$

则它们之间有关系

$$a_{k_1k_2\cdots k_s}^{l_1l_2\cdots l_t}=\sum_{i_1,\cdots,i_s=1}^{n}\sum_{j_1,\cdots,j_t=1}^{n}b_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t}q_{k_1i_1}q_{k_2i_2}\cdots q_{k_si_s}p_{l_1j_1}p_{l_2j_2}\cdots p_{l_tj_t},$$

$$1\leqslant k_1,\cdots,k_s,l_1,\cdots,l_t\leqslant n,\ \text{又}Q=(q_{ij})=(P^{-1})'.$$

证 由定理 8.1.3，记$\{\alpha_1,\cdots,\alpha_n\}$，$\{\beta_1,\cdots\beta_n\}$之对偶基分别为$\{f_1,\cdots,f_n\}$，$\{h_1,\cdots,h_n\}$，则由$\beta_i=\sum\limits_{j=1}^{n}p_{ji}\alpha_j$，$1\leqslant i\leqslant n$ 便有$h_i=\sum\limits_{j=1}^{n}q_{ji}f_j$，$1\leqslant i\leqslant n$. 另一方面，

$$f=\sum_{k_1,\cdots,k_s=1}^{n}\sum_{l_1,\cdots,l_t=1}^{n}a_{k_1k_2\cdots k_s}^{l_1l_2\cdots l_t}f_{k_1}\otimes\cdots\otimes f_{k_s}\otimes\alpha_{l_1}\otimes\cdots\otimes\alpha_{l_t},$$

又

$$\begin{aligned}f&=\sum_{i_1,\cdots,i_s=1}^{n}\sum_{j_1,\cdots,j_t=1}^{n}b_{i_1\cdots i_s}^{j_1\cdots j_t}h_{i_1}\otimes\cdots\otimes h_{i_s}\otimes\beta_{j_1}\otimes\cdots\otimes\beta_{j_t}.\\&=\sum_{i_1,\cdots,i_s=1}^{n}\sum_{j_1,\cdots,j_t=1}^{n}b_{i_1\cdots i_s}^{j_1\cdots j_t}\left(\sum_{k_1=1}^{n}q_{k_1i_1}f_{k_1}\right)\otimes\cdots\otimes\left(\sum_{k_s=1}^{n}q_{k_si_s}f_{k_s}\right)\\&\quad\otimes\left(\sum_{l_1=1}^{n}p_{l_1j_1}\alpha_{l_1}\right)\otimes\cdots\otimes\left(\sum_{l_t=1}^{n}p_{l_tj_t}\alpha_{l_t}\right)\\&=\sum_{k_1,\cdots,k_s=1}^{n}\sum_{l_1,\cdots,l_t=1}^{n}\left(\sum_{i_1,\cdots,i_s=1}^{n}\sum_{j_1,\cdots,j_t=1}^{n}b_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t}q_{k_1i_1}\cdots q_{k_si_s}p_{l_1j_1}\cdots p_{l_tj_t}\right)\\&\quad\times f_{k_1}\otimes\cdots\otimes f_{k_s}\otimes\alpha_{l_1}\otimes\cdots\otimes\alpha_{l_t}.\end{aligned}$$

这证明了定理. 证完.

由于$\mathscr{D}_0^s$为s阶共变张量构成的线性空间. 于是共变张量全体构成$\mathscr{D}$的子空间

$$\mathscr{D}_0=\sum_{s=0}^{\infty}\mathscr{D}_0^s=\sum_{s=0}^{\infty}(\mathscr{D}_0^1)^s=\sum_{s=0}^{\infty}(\mathfrak{L}^*)^s.$$

显然$\mathscr{D}_0$为张量代数$\mathscr{D}$的子代数.

定义 n 维线性空间 $\mathfrak{L}$ 上 s 阶共变张量 f 称为**对称共变张量**，如果对 $1,2,\cdots,s$ 的任意排列 $i_1i_2\cdots i_s$，有

$$f(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_s})=f(\xi_1,\xi_2,\cdots,\xi_s).$$

s 阶共变张量 f 称为**斜对称的**，如果

$$f(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_s})=\delta^{12\cdots s}_{i_1i_2\cdots i_s}f(\xi_1,\xi_2,\cdots,\xi_s).$$

定理 8.2.4 在 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 记 $\{f_1,f_2,\cdots,f_n\}$ 为它在对偶空间 $\mathfrak{L}^*$ 中的对偶基. 则 $\mathfrak{L}$ 上任一 s 阶对称共变张量

$$f=\sum_{1\leqslant i_1\leqslant\cdots\leqslant i_s\leqslant n}a_{i_1i_2\cdots i_s}\sum_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}f_{j_1}\otimes\cdots\otimes f_{j_s}.$$

这证明了 s 阶对称共变张量全体构成线性空间，有基

$$\sum_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}f_{j_1}\otimes\cdots\otimes f_{j_s},\quad 1\leqslant i_1\leqslant i_2\leqslant\cdots\leqslant i_s\leqslant n,$$

其中约定符号

$$\begin{pmatrix}i_1i_2\cdots i_s\\ j_1j_2\cdots j_s\end{pmatrix}$$

的意义为：当 $i_1,i_2,\cdots,i_s$ 为 s 个不同数时，$j_1j_2\cdots j_s$ 为 $i_1,i_2,\cdots,i_s$ 的排列；当 $i_1,i_2,\cdots,i_s$ 中有数值相同时，我们仍看作不同对象，且在此意义下 $j_1j_2\cdots j_s$ 为 $i_1,i_2,\cdots,i_s$ 的排列. 所以和号总共有 $s!$ 项.

证 今

$$f=\sum_{i_1,\cdots,i_s=1}^{n}a_{i_1i_2\cdots i_s}f_{i_1}\otimes\cdots\otimes f_{i_s},$$

其中

$$a_{i_1i_2\cdots i_s}=f(\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_s}).$$

于是视 $i_1,i_2,\cdots,i_s$ 为 s 个不同符号，取它的任意排列 $j_1j_2\cdots j_s$，则有

$$f(\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_s})=f(\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_s}).$$

这证明了

$$a_{i_1i_2\cdots i_s}=a_{j_1j_2\cdots j_s}.$$

所以证明了

$$f=\sum_{1\leqslant i_1\leqslant\cdots\leqslant i_s\leqslant n}a_{i_1i_2\cdots i_s}\sum_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}f_{j_1}\otimes f_{j_2}\otimes\cdots\otimes f_{j_s}.$$

反之，显然

$$\sum_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}f_{j_1}\otimes\cdots\otimes f_{j_s}$$

为对称共变张量. 而且当 $1\leqslant i_1\leqslant\cdots\leqslant i_s\leqslant n$ 时，$\sum\limits_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}f_{j_1}\otimes\cdots\otimes f_{j_s}$ 线性无关. 这证明了它们是一组基. 证完.

最重要的张量是斜对称共变张量.

定理 8.2.5 在 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 记 $\{f_1,f_2,\cdots,f_n\}$ 为它在对偶空间 $\mathfrak{L}^*$ 中的对偶基，则 $\mathfrak{L}$ 上任一 s 阶斜对称共变张量

$$f=\sum_{1\leqslant i_1<\cdots<i_s\leqslant n}a_{i_1i_2\cdots i_s}\sum_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}\delta^{i_1i_2\cdots i_s}_{j_1j_2\cdots j_s}f_{j_1}\otimes\cdots\otimes f_{j_s}.$$

所以所有 s 阶斜对称共变张量构成 $\binom{n}{s}$ 维线性空间 Λ^s，它有基

$$F_{i_1i_2\cdots i_s}=\frac{1}{s!}\sum_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}\delta^{i_1i_2\cdots i_s}_{j_1j_2\cdots j_s}f_{j_1}\otimes\cdots\otimes f_{j_s},\quad 1\leqslant i_1<\cdots<i_s\leqslant n.$$

特别：$\Lambda^0=\mathfrak{D}^0_0$，$\Lambda^1=\mathfrak{L}^*$，$\Lambda^{n+1}=\Lambda^{n+2}=\cdots=0$，因此张量代数 $\varnothing$ 中所有斜对称共变张量构成 2^n 维线性空间：

$$\Lambda=\Lambda^0\oplus\Lambda^1\oplus\cdots\oplus\Lambda^n.$$

证 今 s 阶共变张量

$$f=\sum_{i_1,i_2,\cdots,i_s=1}^{n} a_{i_1i_2\cdots i_s} f_{i_1}\otimes f_{i_2}\otimes\cdots\otimes f_{i_s},$$

其中

$$a_{i_1i_2\cdots i_s}=f(\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_s}).$$

由于 f 为斜对称的，这等价于

$$a_{j_1j_2\cdots j_s}=\delta^{i_1i_2\cdots i_s}_{j_1j_2\cdots j_s}a_{i_1i_2\cdots i_s},$$

其中 $j_1j_2\cdots j_s$ 为 s 个符号 $i_1,i_2,\cdots,i_s$ 的排列.

当 $i_1,i_2,\cdots,i_s$ 中有两个不同指标对应之数 $i_j=i_k$. 于是

$$a_{i_1\cdots i_{j-1}i_ji_{j+1}\cdots i_{k-1}i_ki_{k+1}\cdots i_s}=-a_{i_1\cdots i_{j-1}i_ki_{j+1}\cdots i_{k-1}i_ji_{k+1}\cdots i_s}$$

这证明了 $a_{i_1i_2\cdots i_s}=0$. 所以

$$\begin{aligned}f&=\sum_{i_1,i_2,\cdots,i_s=1}^{n} a_{i_1i_2\cdots i_s}f_{i_1}\otimes\cdots\otimes f_{i_s}\\&=\sum_{\substack{i_1,i_2,\cdots,i_s=1\\ i_1,i_2,\cdots,i_s\text{两两不等}}}^{n} a_{i_1i_2\cdots i_s}f_{i_1}\otimes\cdots\otimes f_{i_s}\\&=\sum_{1\leqslant i_1<\cdots<i_s\leqslant n}a_{i_1i_2\cdots i_s}\sum_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}\delta^{i_1i_2\cdots i_s}_{j_1j_2\cdots j_s}f_{j_1}\otimes f_{j_2}\otimes\cdots\otimes f_{j_s}\end{aligned}$$

显然 $F_{i_1i_2\cdots i_s}=\dfrac{1}{s!}\sum\limits_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}\delta^{i_1i_2\cdots i_s}_{j_1j_2\cdots j_s}f_{j_1}\otimes f_{j_2}\otimes\cdots\otimes f_{j_s}$ 为 s 阶斜对称共变张量，且取遍 $1\leqslant i_1<\cdots<i_s\leqslant n$ 时，它们线性无关. 这证明了 Λ^s 为 $\binom{n}{s}$ 维线性子空间. 证完.

抽象的行列式定义，可以用斜对称共变张量来叙述.

定义 记 Ω 为 $n\times 1$ 矩阵全体构成的线性空间. 则 Ω 上 n 阶斜对称共变张量

$$s!F_{12\cdots n}(\xi_1,\xi_2,\cdots,\xi_n)$$

称为由 n 个列向量 $\xi_1,\cdots,\xi_n$ 决定的**行列式**.

为了证明上面行列式定义和第二章的定义等价. 我们可以更

一般地给出

定理 8.2.6 在 n 维线性空间 $\mathfrak{L}$ 中取基 $\{\alpha_1, \cdots, \alpha_n\}$. 记 $\{f_1, \cdots, f_n\}$ 为它在对偶空间 $\mathfrak{L}^*$ 中的对偶基. 记变向量

$$\xi_i = \sum_{j=1}^{n} x_{ij}\alpha_j, \qquad i=1,2,\cdots,s,$$

则 $\mathfrak{L}$ 上任一 s 阶斜对称共变张量

$$f = \sum_{1\leqslant i_1<\cdots<i_s\leqslant n} a_{i_1 i_2\cdots i_s} F_{i_1 i_2\cdots i_s},$$

其中

$$F_{i_1 i_2\cdots i_s}(\xi_1, \cdots, \xi_s) = \frac{1}{s!}\det\begin{pmatrix} x_{1i_1} & \cdots & x_{1i_s} \\ \vdots & & \vdots \\ x_{si_1} & \cdots & x_{si_s}\end{pmatrix},$$

$$1\leqslant i_1<i_2<\cdots<i_s\leqslant n.$$

特别

$$F_{12\cdots n}(\xi_1, \cdots, \xi_n) = \frac{1}{n!}\det\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn}\end{pmatrix}.$$

证 今

$$\begin{aligned} F_{i_1 i_2\cdots i_s}(\xi_1, \cdots, \xi_s) &= \frac{1}{s!}\sum_{\binom{i_1 i_2\cdots i_s}{j_1 j_2\cdots j_s}} \delta^{i_1 i_2\cdots i_s}_{j_1 j_2\cdots j_s}(f_{j_1}\otimes\cdots\otimes f_{j_s})(\xi_1, \cdots \xi_s) \\ &= \frac{1}{s!}\sum_{\binom{i_1 i_2\cdots i_s}{j_1 j_2\cdots j_s}} \delta^{i_1 i_2\cdots i_s}_{j_1 j_2\cdots j_s} f_{j_1}(\xi_1)\cdots f_{j_s}(\xi_s) \\ &= \frac{1}{s!}\sum_{\binom{i_1 i_2\cdots i_s}{j_1 j_2\cdots j_s}} \delta^{i_1 i_2\cdots i_s}_{j_1 j_2\cdots j_s} x_{1j_1}x_{2j_2}\cdots x_{sj_s} \\ &= \frac{1}{s!}\det\begin{pmatrix} x_{1i_1} & \cdots & x_{1i_s} \\ \vdots & & \vdots \\ x_{si_1} & \cdots & x_{si_s}\end{pmatrix}. \end{aligned}$$

证完.

为了在斜对称共变张量之间引进另一种近代数学中重要的乘法——外乘. 我们先给出

引理 8.2.3 设 $1\leqslant s, t\leqslant n$, 且 $f\in\Lambda^s, g\in\Lambda^t$, 则

$$h(\xi_1,\cdots,\xi_{s+t})=\frac{1}{(s+t)!}\sum_{\binom{12\cdots(s+t)}{i_1i_2\cdots i_{s+t}}}\delta^{12\cdots(s+t)}_{i_1i_2\cdots i_{s+t}}$$

$$\times f(\xi_{i_1},\cdots,\xi_{i_s})g(\xi_{i_{s+1}},\cdots,\xi_{i_{s+t}})$$

定义了一个 $s+t$ 阶斜对称共变张量 h, 即 $h\in\Lambda^{s+t}$.

证 显然 $h(\xi_1,\cdots,\xi_{s+t})$ 为 $s+t$ 重线性函数. 下面证它是斜对称的. 今任取 $1,2,\cdots,s+t$ 的排列 $j_1j_2\cdots j_{s+t}$, 我们来证

$$h(\xi_{j_1}\cdots,\xi_{j_{s+t}})=\delta^{12\cdots(s+t)}_{j_1j_2\cdots j_{s+t}}\ h(\xi_1,\cdots,\xi_{s+t})$$

事实上, 当 $i_1i_2\cdots i_{s+t}$ 遍历 $j_1,j_2\cdots,j_{s+t}$ 的所有排列时, 也遍历 $1,2,\cdots,(s+t)$ 的所有排列, 所以

$$\begin{aligned}
&h(\xi_{j_1},\xi_{j_2},\cdots,\xi_{j_{s+t}})\\
&=\frac{1}{(s+t)!}\sum_{\binom{j_1j_2\cdots j_{s+t}}{i_1i_2\cdots i_{s+t}}}\delta^{j_1j_2\cdots j_{s+t}}_{i_1i_2\cdots i_{s+t}}f(\xi_{i_1},\cdots,\xi_{i_s})g(\xi_{i_{s+1}},\cdots,\xi_{i_{s+t}})\\
&=\frac{1}{(s+t)!}\delta^{12\cdots(s+t)}_{j_1j_2\cdots j_{s+t}}\sum_{\binom{12\cdots(s+t)}{i_1i_2\cdots i_{s+t}}}\delta^{12\cdots(s+t)}_{i_1i_2\cdots i_{s+t}}f(\xi_{i_1},\cdots\xi_{i_s})\\
&\qquad\times g(\xi_{i_{s+1}},\cdots,\xi_{i_{s+t}})\\
&=\delta^{12\cdots(s+t)}_{j_1j_2\cdots j_{s+t}}h(\xi_1,\xi_2,\cdots,\xi_{s+t})
\end{aligned}$$

证完.

定义 记 Λ 为 Ω 上所有斜对称共变张量场线性生成的子空间. 在 Λ 中定义外乘"$\wedge$"如下:

(1) $\sum\limits_{i=0}^{n}f_i\wedge\sum\limits_{j=0}^{n}g_j=\sum\limits_{i,j=0}^{n}f_i\wedge g_j, \forall f_i,\quad g_i\in\Lambda^i,\quad 0\leqslant i\leqslant n;$

(2) $f_0\wedge g_i=f_0g_i,\quad f_i\wedge g_0=g_0f_i,\quad \forall$

$$f_i, g_i \in \Lambda^i, \quad 0 \leqslant i \leqslant n;$$

(3) 当 $1 \leqslant s, t \leqslant n$, $f \in \Lambda^s$, $g \in \Lambda^t$. 则由引理 8.2.3 定义的 $h \in \Lambda^{s+t}$ 有 $f \wedge g = h$. 换句话说

$$(f \wedge g)(\xi_1, \cdots, \xi_{s+t}) = \frac{1}{(s+t)!} \sum_{\binom{12\cdots(s+t)}{i_1 i_2 \cdots i_{s+t}}} \delta^{12\cdots(s+t)}_{i_1 i_2 \cdots i_{s+t}}$$

$$\cdot f(\xi_{i_1}, \cdots, \xi_{i_s}) g(\xi_{i_{s+1}}, \cdots, \xi_{i_{s+t}})$$

由定义，我们有

定理 8.2.7 在 2^n 维线性空间 $\Lambda = \sum_{i=0}^{n} \Lambda^i$ 中引进外乘"$\wedge$"后，Λ 构成一个 2^n 维结合代数，称为外代数或称为 Grassmann 代数. 它有基

$$1, \quad F_{i_1 i_2 \cdots i_s}, \qquad 1 \leqslant i_1 < \cdots < i_s \leqslant n, \quad 1 \leqslant s \leqslant n.$$

且有

$$F_{i_1 i_2 \cdots i_s} = f_{i_1} \wedge f_{i_2} \wedge \cdots \wedge f_{i_s},$$

其中 $\{f_1, \cdots, f_n\}$ 为 $\mathfrak{L}$ 的基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 在对偶空间 $\mathfrak{L}^*$ 中的对偶基.

证 任取 $f \in \mathfrak{L}^* = \Lambda^1$，则

$$(f \wedge f)(\xi_1, \xi_2) = \frac{1}{2!} \sum_{\binom{1\ 2}{i_1 i_2}} \delta^{12}_{i_1 i_2} f(\xi_{i_1}) f(\xi_{i_2})$$

$$= \frac{1}{2} (f(\xi_1) f(\xi_2) - f(\xi_2) f(\xi_1)) = 0$$

所以

$$f \wedge f = 0.$$

令 $F_{i_1} = f_{i_1}$. 设 $F_{i_1 i_2 \cdots i_{s-1}} = f_{i_1} \wedge f_{i_2} \wedge \cdots \wedge f_{i_{s-1}}$. 则

$$(f_{i_1} \wedge f_{i_2} \wedge \cdots \wedge f_{i_s})(\xi_1, \xi_2, \cdots, \xi_s)$$

$$= (F_{i_1 i_2 \cdots i_{s-1}} \wedge f_{i_s}) \qquad (\xi_1, \xi_2, \cdots, \xi_s)$$

由引理 8.2.3,

$$(f_{i_1}\wedge f_{i_2}\wedge\cdots\wedge f_{i_s})(\xi_1,\xi_2,\cdots,\xi_s)$$

$$=\frac{1}{s!}\sum_{\binom{12\cdots s}{j_1j_2\cdots j_s}}\delta^{12\cdots s}_{j_1j_2\cdots j_s}F_{i_1i_2\cdots i_{s-1}}(\xi_{j_1},\cdots\xi_{j_{s-1}})f_{i_s}(\xi_{j_s}).$$

记

$$\xi_i=\sum_{j=1}^{n}x_{ij}\alpha_j,\quad 1\leqslant i\leqslant s,$$

则由定理 8.2.6,有

$$(f_{i_1}\wedge f_{i_2}\wedge\cdots\wedge f_{i_s})(\xi_1,\xi_2,\cdots,\xi_s)$$

$$=\frac{1}{s!}\sum_{\binom{12\cdots s}{j_1j_2\cdots j_s}}\delta^{12\cdots s}_{j_1j_2\cdots j_s}\frac{1}{(s-1)!}\det\begin{pmatrix}x_{j_1i_1} & \cdots & x_{j_1i_{s-1}}\\ \vdots & & \vdots\\ x_{j_{s-1}i_1} & \cdots & x_{j_{s-1}i_{s-1}}\end{pmatrix}x_{j_si_s}$$

$$=\frac{1}{(s-1)!s!}\sum_{\binom{12\cdots s}{j_1j_2\cdots j_s}}\sum_{\binom{j_1j_2\cdots j_{s-1}}{k_1k_2\cdots k_{s-1}}}\delta^{12\cdots s}_{j_1j_2\cdots j_s}\delta^{j_1j_2\cdots j_{s-1}}_{k_1k_2\cdots k_{s-1}}$$

$$\times x_{k_1i_1}x_{k_2i_2}\cdots x_{k_{s-1}i_{s-1}}x_{j_si_s}$$

现在将$k_1k_2\cdots k_{s-1}$作一系列对换变为$j_1j_2\cdots j_{s-1}$，它自然将$k_1k_2\cdots k_{s-1}j_s$变为$j_1j_2\cdots j_{s-1}j_s$. 所以

$$\delta^{j_1j_2\cdots j_{s-1}}_{k_1k_2\cdots k_{s-1}}=\delta^{j_1j_2\cdots j_{s-1}j_s}_{k_1k_2\cdots k_{s-1}j_s}.$$

因此

$$(f_{i_1}\wedge\cdots\wedge f_{i_s})(\xi_1,\cdots,\xi_s)$$

$$=\frac{1}{(s-1)!s!}\sum_{\binom{12\cdots s}{j_1j_2\cdots j_s}}\sum_{\binom{j_1j_2\cdots j_{s-1}j_s}{k_1k_2\cdots k_{s-1}j_s}}\delta^{12\cdots(s-1)s}_{k_1k_2\cdots k_{s-1}j_s}x_{k_1i_1}\cdots x_{k_{s-1}i_{s-1}}x_{j_si_{s-1}}$$

$$=\frac{1}{s!}\sum_{\binom{12\cdots s}{j_1j_2\cdots j_s}}\delta^{12\cdots s}_{j_1j_2\cdots j_s}x_{j_1i_1}x_{j_2j_2}\cdots x_{j_si_s}$$

$$=F_{i_1i_2\cdots i_s}(\xi_1,\cdots,\xi_s).$$

证完.

习题 8.2

1. 试证：外代数 $\Lambda=\sum_{j=0}^{n}\Lambda^j$ 有性质

$$f\wedge g=(-1)^{st}g\wedge f,\quad \forall f\in\Lambda^s,\quad g\in\Lambda^t,\quad 0\leqslant s,t\leqslant n.$$

2. 记 $C^1(x,y)$ 为两个变量 x，y 的一切连续可微函数全体构成的集合。试证：偏微分算子

$$X=f(x,y)\frac{\partial}{\partial x}+g(x,y)\frac{\partial}{\partial y},\quad \forall f,g\in C^1(x,y)$$

全体构成的实线性空间 $\mathscr{L}$. 记

$$u(x,y)dx+v(x,y)dy,\quad \forall u,v\in C^1(x,y)$$

为 $\mathscr{L}$ 上实线性函数，定义为

$$dx\left(\frac{\partial}{\partial x}\right)=1,\quad dx\left(\frac{\partial}{\partial y}\right)=dy\left(\frac{\partial}{\partial x}\right)=0,\quad dy\left(\frac{\partial}{\partial y}\right)=1,$$

又

$$(udx+vdy)\left(f\frac{\partial}{\partial x}+g\frac{\partial}{\partial y}\right)=uf+vg.$$

引进外积

$$(udx+vdy)\wedge(pdx+qdy)=\det\begin{pmatrix}u & v\\ p & q\end{pmatrix}\cdot dx\wedge dy.$$

于是

$$dx\wedge dx=0,\quad dx\wedge dy=-dy\wedge dx,\quad dy\wedge dy=0.$$

试证：对变数变换

$$\tilde{x}=f(x,y),\quad \tilde{y}=g(x,y),\quad f,g\in C^1(x,y)$$

记 Jacobian 为

$$J(x,y)=\det\begin{pmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y}\end{pmatrix},$$

则有

$$d\tilde{x}\wedge d\tilde{y}=J(x,y)dx\wedge dy$$

由此可知二重积分应记作

$$\iint_D F(x,y)dx\wedge dy.$$

3. 设 $f(\xi_1,\cdots,\xi_s)$ 为 Ω 上 s 重对称线性函数．令

$$x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$$

则 $f(x,x,\cdots;x)$ 为 $x_1, x_2, \cdots, x_n$ 的 s 次齐次多项式．反之，任给 s 次齐次多项式 $F(x_1,\cdots,x_n)=F(x)$，试证：唯一存在一个 Ω 上 s 重对称线性函数 $f(\xi_1\cdots,\xi_s)$，使得 $f(x,\cdots,x)=F(x)$.

第九章 Euclid 空 间

§ 9.1 Euclid 空 间

在这一章，我们限制在实数范围内进行讨论，即我们取纯量为实数.

定义 n 维实线性空间 $\mathfrak{L}$ 上定正对称双线性函数 称为 $\mathfrak{L}$上的**内积**.

在 n 维实线性空间 $\mathfrak{L}$ 上取定内积，记作(α,β)则有

引理 9.1.1 (Cauchy不等式) 在 n 维实线性空间 $\mathfrak{L}$ 上取定内积(α,β)，则有

$$(\alpha,\beta)^2 \leqslant (\alpha,\alpha)(\beta,\beta), \quad \forall \alpha,\beta \in \mathfrak{L}$$

且等式成立当且仅当 α 和β 线性相关.

证 引进实参数 t，则有 $\alpha + t\beta \in \mathfrak{L}$. 于是

$$(\alpha+t\beta,\alpha+t\beta) = (\alpha,\alpha) + 2t(\alpha,\beta) + t^2(\beta,\beta) \geqslant 0,$$
$$\forall t \in \mathbf{R}.$$

当 $\beta=0$，不等式$(\alpha,\beta)^2 \leqslant (\alpha,\alpha)(\beta,\beta)$自然成立. 当 $\beta \neq 0$，于是$(\beta,\beta) > 0$. 因此二次多项式的判别式$\leqslant 0$，即有

$$(2(\alpha,\beta))^2 - 4(\alpha,\alpha)(\beta,\beta) \leqslant 0$$

此即 Cauchy 不等式成立. 等号成立当且仅当二次 多项式有实根 t_0，即有$(\alpha+t_0\beta,\alpha+t_0\beta)=0$. 由于内积为定正对称双线性函数，所以有 $\alpha + t_0\beta = 0$，此即 α,β 线性相关. 证完.

定义 在 n 维实线性空间 $\mathfrak{L}$ 上取定内积(α,β). 则数值

$$|\alpha| = \sqrt{(\alpha,\alpha)} \geqslant 0$$

称为向量 α 的**长度**. 长度为 1 的向量称为**单位向量**.

显然,长度为零的向量只有零向量.

定义 在 n 维实线性空间 $\mathfrak{L}$ 上取定内积 (α,β). 当 $\alpha=0$ 或 $\beta=0$ 时,我们约定向量 α 和 β 之夹角为直角. 当 $\alpha\neq 0,\beta\neq 0$ 时,数值

$$\theta=\cos^{-1}\frac{(\alpha,\beta)}{|\alpha||\beta|}$$

称为向量 α 和 β 的**夹角**. 这里约定 $\cos^{-1}$ 取主值,即有

$$0\leqslant\theta\leqslant\pi.$$

夹角为直角时,即当 $(\alpha,\beta)=0$ 时,我们称向量 α 和 β **互相正交**,记作 $\alpha\perp\beta$.

显然,只有零向量才和任何向量正交.

定义 在 n 维实线性空间 $\mathfrak{L}$ 上取定内积 (α,β), 则 $\mathfrak{L}$ 称为关于内积 (α,β) 的 Euclid 空间或称为**欧氏空间**.

定义 记 $\mathfrak{L}$ 为 n 维 Euclid 空间, 内积为 (α,β). $\mathfrak{L}$ 中基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 称为**标准正交基**,如果

$$(\alpha_i,\alpha_j)=\delta_{ij},\quad 1\leqslant i,j\leqslant n,$$

即 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 n 个两两正交的单位向量.

上面给出标准正交基的存在性.

引理 9.1.2 设 $\beta_1,\cdots,\beta_r$ 为 n 维 Euclid 空间 $\mathfrak{L}$ 中 r 个两两正交的单位向量,则它们线性无关,所以 $r\leqslant n$. 特别 $\mathfrak{L}$ 中 n 个两两正交的单位向量必为标准正交基.

证 设存在常数 $a_1,\cdots,a_r$, 使得 $\sum\limits_{j=1}^{r}a_j\beta_j=0$. 于是

$$\begin{aligned}0=(0,\beta_i)&=\left(\sum_{j=1}^{r}a_j\beta_j,\ \beta_i\right)\\&=\sum_{j=1}^{r}a_j(\beta_j,\beta_i)=a_i,\quad i=1,2,\cdots,r.\end{aligned}$$

这证明了 $\beta_1,\beta_2,\cdots,\beta_r$ 线性无关. 证完.

定理 9.1.1 (Schmidt 正交化) 设 $\mathfrak{L}$ 为 n 维 Euclid 空间. 在 $\mathfrak{L}$ 中任取一组基 $\{\beta_1,\beta_2,\cdots,\beta_n\}$, 则存在实数 b_{ij}, $1\leqslant j\leqslant i\leqslant n$, 使得 $b_{11}>0,\cdots,b_{nn}>0$, 且

$$\alpha_1=b_{11}\beta_1,$$
$$\alpha_2=b_{21}\beta_1+b_{22}\beta_2,$$
$$\cdots\cdots$$
$$\alpha_n=b_{n1}\beta_1+b_{n2}\beta_2+\cdots+b_{nn}\beta_n$$

为 $\mathfrak{L}$ 中的标准正交基.

证 下面用归纳法给出构造 b_{ij}, $1\leqslant j\leqslant i\leqslant n$ 的方法. 这个方法称为 Schmidt 正交化.

令 $\beta_1\neq 0$. 将它单位化, 即取

$$\alpha_1=\frac{1}{|\beta_1|}\beta_1=b_{11}\beta_1,$$

所以 $b_{11}>0$, 且 α_1 为单位向量. 令

$$\bar{\alpha}_2=\beta_2+a_{12}\alpha_1=\beta_2+a_{21}b_{11}\beta_1\neq 0,$$

使得 $\bar{\alpha}_2\perp\alpha_1$, 即有

$$0=(\bar{\alpha}_2,\alpha_1)=(\beta_2,\alpha_1)+a_{21}(\alpha_1,\alpha_1)=(\beta_2,\alpha_1)+a_{21}$$

于是定出

$$a_{21}=-(\beta_2,\alpha_1),$$

即取

$$\bar{\alpha}_2=\beta_2-(\beta_2,\alpha_1)\alpha_1=\beta_2-b_{11}{}^2(\beta_2,\beta_1)\beta_1,$$

再将 $\bar{\alpha}_2$ 单位化, 即取

$$\alpha_2=|\bar{\alpha}_2|^{-1}\bar{\alpha}_2=b_{21}\beta_1+b_{22}\beta_2$$

所以 $b_{22}=|\bar{\alpha}_2|^{-1}>0$, 且 α_1,α_2 为两两正交的单位向量.

由归纳法假设, 所以存在 r 个两两正交的单位向量

$$\alpha_1=b_{11}\beta_1,$$

$$\alpha_2=b_{21}\beta_1+b_{22}\beta_2,$$

$$\cdots\cdots$$

$$\alpha_r=b_{r1}\beta_1+b_{r2}\beta_2+\cdots+b_{rr}\beta_r,$$

其中 $b_{11}>0,\cdots,b_{rr}>0$. 取

$$\begin{aligned}\tilde{\alpha}_{r+1}&=\beta_{r+1}+a_{r+1,1}\alpha_1+\cdots+a_{r+1,r}\alpha_r\\&=\beta_{r+1}+\tilde{a}_{r+1,1}\beta_1+\cdots+\tilde{a}_{r+1,r}\beta_r\neq 0.\end{aligned}$$

使得 $(\tilde{\alpha}_{r+1},\alpha_j)=0,\ j=1,2,\cdots,r$. 即

$$\begin{aligned}0=(\tilde{\alpha}_{r+1},\alpha_j)&=(\beta_{r+1},\alpha_j)+a_{r+1,1}(\alpha_1,\alpha_j)+\cdots\\&+a_{r+1,r}(\alpha_r,\alpha_j)=(\beta_{r+1},\alpha_j)+a_{r+1,j},\end{aligned}$$

所以取

$$a_{r+1,j}=-(\beta_{r+1},\alpha_j),\quad j=1,2,\cdots,r$$

便有 $\alpha_j\perp\tilde{\alpha}_{r+1}$, $1\leqslant j\leqslant r$. 即

$$\tilde{\alpha}_{r+1}=\beta_{r+1}-(\beta_{r+1},\alpha_1)\alpha_1-\cdots-(\beta_{r+1},\alpha_r)\alpha_r.$$

由于 $\tilde{\alpha}_{r+1}\neq 0$, 将它单位化,即取

$$\alpha_{r+1}=|\tilde{\alpha}_{r+1}|^{-1}\tilde{\alpha}_{r+1}$$

则$\alpha_1,\alpha_2,\cdots\alpha_r,\alpha_{r+1}$为两两正交的单位向量,且

$$\alpha_{r+1}=b_{r+1,1}\beta_1+\cdots+b_{r+1,r}\beta_r+b_{r+1,r+1}\beta_{r+1},$$

其中 $b_{r+1,r+1}>0$. 由归纳法,作了第 n步. 由引理 9.1.2, 便得到适合定理条件的标准正交基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 证完.

定理 9.1.2 在n 维 Euclid 空间 $\mathfrak{L}$ 中任取 r 个两两正交的单位向量 $\alpha_1,\alpha_2,\cdots,\alpha_r$, 于是 $r\leqslant n$. 且在 $\mathfrak{L}$ 中存在 $n-r$ 个向量 $\alpha_{r+1},\cdots,\alpha_n$, 使得 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$ 为 $\mathfrak{L}$ 的标准正交基.

证 由第六章可知: 在 $\mathfrak{L}$ 中存在 $n-r$ 个向量 $\beta_{r+1},\cdots,\beta_n$, 使得 $\alpha_1,\cdots,\alpha_r,\beta_{r+1},\cdots,\beta_n$ 为 $\mathfrak{L}$ 的基. 对这组基作 Schmidt 正交化, 所以得到 $\mathfrak{L}$ 的标准正交基 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$. 证完.

最常用的性质为 $r=1$ 的情形, 即我们有

定理 9.1.3 在 n 维 Euclid 空间 $\mathfrak{L}$ 中任取单位向量 α, 则存

在一组标准正交基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，使得 α 为指定的第 j 个基向量 α_j，特别，单位向量能成为一组标准正交基的第一个向量.

标准正交基的特点为：

定理 9.1.4 设 Ω 为关于内积 (α,β) 的 n 维 Euclid 空间. $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 为 Ω 的标准正交基. 记

$$\alpha=\sum_{j=1}^{n}x_j\alpha_j,\quad \beta=\sum_{j=1}^{n}y_j\alpha_j,$$

则有

$$(\alpha,\beta)=\sum_{j=1}^{n}x_jy_j.$$

即在标准正交基下，内积的方阵表示为单位方阵.

证 由 $(\alpha_i,\alpha_j)=\delta_{ij}$，于是

$$(\alpha,\beta)=\left(\sum_{i=1}^{n}x_i\alpha_i,\ \sum_{j=1}^{n}y_j\alpha_j\right)$$

$$=\sum_{i,j=1}^{n}x_iy_j(\alpha_i,\alpha_j)=\sum_{j=1}^{n}x_jy_j,$$

证完.

定义 n 阶实方阵 O 称为**实正交方阵**，如果

$$OO'=O'O=I^{(n)},$$

即 $O'=O^{-1}$. 所有 n 阶实正交方阵构成的集合记作 $O(n,R)$，或简记为 $O(n)$，称为 **n 阶实正交群**.

引理 9.1.3 n 阶实正交群 $O(n)$ 有性质

(1) 任取 $O_1,O_2\in O(n)$，则 $O_1O_2\in O(n)$，

(2) 任取 $O\in O(n)$，则 $O^{-1}=O'\in O(n)$，

(3) n 阶单位方阵 $I\in O(n)$，

(4) 任取 $O\in O(n)$，则 $\det O=\pm1$.

证 今由 O_1，$O_2\in O(n)$，即有 $O_1O_1'=O_1'O_1=I$，$O_2O_2'=$

$O_2'O_2=I$. 于是

$$(O_1O_2)(O_1O_2)'=O_1O_2O_2'O_1'=O_1O_1'=I,$$

$$(O_1O_2)'(O_1O_2)=O_2'O_1'O_1O_2=O_2'O_2=I.$$

这证明了 $O_1O_2\in O(n)$，再任取 $O\in O(n)$，于是 $O^{-1}=O'$，因此只要证 $O'\in O(n)$即可，今 $O'(O')'=O'O=I$，$(O')'O'=OO'=I$，所以 $O'\in O(n)$，因此 $O^{-1}\in O(n)$. 由实正交方阵定义可知 n 阶单位方阵在 $O(n)$ 中. 最后，由$OO'=I$，双方取行列式，所以有 $1=\det OO'=\det O\det O'=(\det O)^2$，即有 $\det O=\pm1$. 证完.

定理 9.1.5 设 $\mathfrak{R}$ 为 n 维 Euclid 空间，内积为(α,β). 任取标准正交基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 再在$\mathfrak{R}$中任取一组基$\{\beta_1,\beta_2,\cdots,\beta_n\}$，使得基变换公式为

$$\beta_i=\sum_{j=1}^{n}a_{ji}\alpha_j,\quad i=1,2,\cdots,n.$$

它对应的非异实方阵

$$A=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}.$$

则$\{\beta_1,\cdots,\beta_n\}$为标准正交基当且仅当$A$为实正交方阵. 所以在 n 维 Euclid 空间 $\mathfrak{R}$ 中取定标准正交基后，实正交方阵和 $\mathfrak{R}$ 中标准正交基间便有一个自然的一一对应关系.

证 令 $\mathfrak{R}$ 中基$\{\beta_1,\beta_2,\cdots,\beta_n\}$ 为标准正交基，即$(\beta_i,\beta_j)=\delta_{ij}$，$i,j=1,2,\cdots,n$. 此即

$$\delta_{ij}=\Big(\sum_{k=1}^{n}a_{ki}\alpha_k,\ \sum_{l=1}^{n}a_{lj}\alpha_l\Big)$$

$$=\sum_{k,l=1}^{n}a_{ki}a_{lj}(\alpha_k,\alpha_l)$$

由于已知 $\{\alpha_1,\cdots,\alpha_n\}$ 为标准正交基，即有 $(\alpha_k,\alpha_l)=\delta_{kl}$，$1\leqslant k$,

$l\leqslant n$，所以证明了$\{\beta_1,\cdots,\beta_n\}$为标准正交基当且仅当

$$\sum_{k=1}^{n}a_{ki}a_{kj}=\delta_{ij},\quad 1\leqslant i,j\leqslant n.$$

此即$A'A=I^{(n)}$，所以$A^{-1}=A'$，即A为n阶实正交方阵．证完．

上面引进的实正交方阵还有许多有用的性质．为此，我们考虑所有$n\times 1$实矩阵构成的实线性空间V_n，在V_n中引进标准内积

$$(x,y)=\left(\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\right)=\sum_{j=1}^{n}x_jy_j=x'y.$$

显然它是V_n上的定正对称双线性函数，且在V_n中有标准正交基

$$e_1=\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix},\quad e_2=\begin{pmatrix}0\\1\\0\\ \vdots\\0\end{pmatrix},\quad\cdots,\quad e_n=\begin{pmatrix}0\\ \vdots\\0\\1\end{pmatrix}.$$

设Ω为n维Euclid空间，它有标准正交基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$．于是对$\Omega$中任两向量$\alpha,\beta$，有

$$\alpha=\sum_{j=1}^{n}x_j\alpha_j\longrightarrow\begin{pmatrix}x_1\\x_2\\ \vdots\\x_n\end{pmatrix}=x\in V_n,$$

$$\beta=\sum_{j=1}^{n}y_j\alpha_j\longrightarrow\begin{pmatrix}y_1\\y_2\\ \vdots\\y_n\end{pmatrix}=y\in V_n,$$

且Ω中内积(α,β)有

$$(\alpha,\beta)=\sum_{j=1}^{n}x_jy_j=(x,y).$$

其中(x,y)为实线性空间V_n中的标准内积. 自然$\alpha_i\longrightarrow e_i, i=1, 2,\cdots,n$.

定理 9.1.6 n阶实正交方阵有性质:

(1) O为n阶实正交方阵当且仅当它的n个行向量为V_n中标准正交基; 当且仅当它的n个列向量也为V_n中标准正交基.

(2) 任给$n\times m$实矩阵O_1,其中$m\leqslant n$. 设

$$O_1'O_1=I^{(m)},$$

则存在$n\times(n-m)$实矩阵O_2,使得n阶实方阵

$$O=(O_1\quad O_2)$$

为实正交方阵. 反之, n阶实正交方阵的前m列构成的$n\times m$实矩阵O_1适合$O_1'O_1=I^{(m)}$.

特别, 任一单位列向量能成为一个n阶实正交方阵的指定的列向量,所以能成为一个n阶实正交方阵的第一个列向量.

(3) 任给$m\times n$实阵矩O_1,其中$m<n$ 设

$$O_1O_1'=I^{(m)}.$$

则存在$(n-m)\times m$实矩阵O_2,使得n阶实方阵

$$O=\begin{pmatrix}O_1\\O_2\end{pmatrix}$$

为实正交方阵. 反之,n阶实正交方阵的前m行构成的$m\times n$实矩阵O_1适合$O_1O_1'=I^{(m)}$.

特别, 任一单位行向量能成为一个n阶实正交方阵的指定的行向量,所以能成为一个n阶实正交方阵的第一个行向量.

证 由于O为实正交方阵当且仅当O'为实正交方阵. 所以我们只要对列向量进行讨论即可. 先证(1). 今给定n阶实方阵

$$O=(\alpha_1\alpha_2\cdots\alpha_n),$$

其中$\alpha_1,\cdots,\alpha_n$为n个$n\times 1$实矩阵. 于是

$$O'O=\begin{pmatrix}\alpha_1'\\ \vdots \\ \alpha_n'\end{pmatrix}(\alpha_1\cdots\alpha_n)=\begin{pmatrix}\alpha_1'\alpha_1 & \cdots & \alpha_1'\alpha_n\\ \vdots & & \vdots\\ \alpha_n'\alpha_1 & \cdots & \alpha_n'\alpha_n\end{pmatrix}.$$

因此 $O'O=I^{(n)}$ 当且仅当 $(\alpha_i,\alpha_j)=\alpha_i'\alpha_j=\delta_{ij}$，$1\leqslant i,j\leqslant n$.
这证明了 O 为实正交方阵当且仅当它的 n 个列向量为 V_n 中一组标准正交基，即证明了(1).

下面证(2). 今 O_1 为 $n\times m$ 实矩阵，有 $O_1'O_1=I^{(m)}$. 此即 O_1 的 m 个列向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 为 V_n 中两两正交的单位向量. 由定理 9.1.2，在 V_n 中存在标准正交基 $\alpha_1,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_n$，于是 $O=(\alpha_1\alpha_2\cdots\alpha_m\alpha_{m+1}\cdots\alpha_n)=(O_1O_2)$ 为实正交方阵. 反之显然，这证明了(2). 证完.

定理 9.1.7 任给 n 阶实方阵 A，则存在 n 阶实正交方阵 O_1，O_2，及 n 阶上三角方阵 T_1，n 阶下三角方阵 T_2，使得 T_1 及 T_2 的对角元素都非负，又有

$$A=O_1T_1=T_2O_2.$$

当 A 为 n 阶非异实方阵时，上述分解唯一.

证 记 n 阶实方阵 A 的 n 个列向量依次为 $\beta_1,\beta_2,\cdots,\beta_n$. 我们用下面办法选取极大线性无关部分组：若 $\beta_1=0,\cdots,\beta_{n_1-1}=0$，$\beta_{n_1}\neq 0$，则取出 β_{n_1}. 若 $\beta_{n_1+1},\cdots,\beta_{n_1+n_2-1}$ 为 $\beta_1,\cdots,\beta_{n_1}$ 的线性组合则舍去，若 $\beta_{n_1+n_2}$ 不是 $\beta_1,\cdots,\beta_{n_1}$ 的线性组合，则取出 $\beta_{n_1+n_2}$，所以有 $\beta_{n-1},\beta_{n_1+n_2}$. 这样依次作下去，便选出了极大线性无关部分组 $\gamma_1,\gamma_2,\cdots,\gamma_r$. 在 V_n 中补足 $n-r$ 个向量 $\gamma_{r+1},\cdots,\gamma_n$，便得到 V_n 的一组基 $\gamma_1,\cdots,\gamma_r,\gamma_{r+1},\cdots,\gamma_n$. 对这组基作 Schmidt 正交化，便得标准正交基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，这时有

$$\gamma_1=b_{11}\alpha_1,$$
$$\gamma_2=b_{21}\alpha_1+b_{22}\alpha_2,$$
$$\cdots\cdots$$

$$\gamma_n = b_{n1}\alpha_1 + b_{n2}\alpha_2 + \cdots + b_{nn}\alpha_n,$$

其中 $b_{11}>0, \cdots, b_{nn}>0$.

由于 $\gamma_1, \cdots, \gamma_r$ 为向量组 $\beta_1, \beta_2, \cdots, \beta_n$ 的极大线性无关部分组,且选取方法是依次从 $\beta_1, \beta_2, \cdots, \beta_n$ 选取. 所以

$$\beta_1 = a_{11}\alpha_1,$$

$$\beta_2 = a_{21}\alpha_1 + a_{22}\alpha_2,$$

$$\cdots\cdots$$

$$\beta_n = a_{n1}\alpha_1 + a_{n2}\alpha_2 + \cdots + a_{nn}\alpha_n,$$

其中 $a_{11} \geqslant 0, \cdots, a_{nn} \geqslant 0$, 令

$$O_1 = (\alpha_1\alpha_2\cdots\alpha_n)$$

为实正交方阵,而

$$A = (\beta_1\beta_2\cdots\beta_n) = (\alpha_1\alpha_2\cdots\alpha_n)\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ & a_{22} & \cdots & a_{n2} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix} = O_1T_1,$$

其中 T_1 为上三角方阵,对角元素非负. 这证明了 $A=O_1T_1$.

对 n 阶实方阵 A' 用上面性质,便证明了 $A=T_2O_2$,其中 O_2 为 n 阶实正交方阵,T_2 为下三角方阵,且对角元素非负.

最后证唯一性. 设 A 为 n 阶非异方阵, 且有分解 $A=O_1T_1$. 于是 $\det T_1 \neq 0$, 即 T_1 的对角元素都是正数. 设 $A=O_1T_1=O_3T_3$, 其中 O_1 及 O_3 为 n 阶实正交方阵,T_1 及 T_3 为 n 阶上三角方阵, 对角元素都大于零. 于是 T_3^{-1} 和 $T_1T_3^{-1}$ 也是上三角方阵, 且对角元素大于零. 但是 $T_1T_3^{-1}=O_1'O_3$是实正交方阵,所以证明了 $T_1T_3^{-1}=I$,即 $T_3=T_1$,因此 $O_3=O_1$. 这证明了分解唯一性. 由 $A'=O_2'T_2'$ 的分解唯一性,便证明了$A=T_2O_2$ 的分解唯一性. 定理证完.

例 设N为实正交方阵O的一个 r 阶子式, M为N的代数余子式. 试证:

$$N = M\det O.$$

证 无妨设N为n阶实正交方阵O的前r行，前r列决定的子矩阵的行列式，即记

$$O=\begin{pmatrix} O_{11}^{(r)} & O_{12} \\ O_{21} & O_{22} \end{pmatrix},$$

其中$\det O_{11}=N$. 于是$\det O_{22}=M$. 所以要证

$$\det O_{11}=\det O_{22}\det O.$$

今$OO'=I$，即

$$O_{11}O_{11}'+O_{12}O_{12}'=I,\quad O_{11}O_{21}'+O_{12}O_{22}'=0,\quad O_{21}O_{21}'+O_{22}O_{22}'=I$$

因此

$$\begin{pmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{pmatrix}\begin{pmatrix} I & O_{21}' \\ 0 & O_{22}' \end{pmatrix}=\begin{pmatrix} O_{11} & 0 \\ O_{21} & I \end{pmatrix}.$$

双方取行列式，便证明了

$$\det O\ \det O_{22}=\det O_{11}.$$

证完.

定义 在n维Euclid空间$\mathfrak{L}$上保持向量长度不变的线性变换称为**正交变换**.

定理 9.1.8 Euclid空间$\mathfrak{L}$上线性变换$\mathscr{A}$为正交变换当且仅当它保持内积不变.

证 设$\mathscr{A}$为正交变换，即有$|\mathscr{A}(\alpha)|=|\alpha|$，$\forall\alpha\in\mathfrak{L}$. 于是$|\mathscr{A}(\alpha+\beta)|=|\alpha+\beta|$，所以

$$(\mathscr{A}(\alpha+\beta),\quad \mathscr{A}(\alpha+\beta))=(\alpha+\beta,\alpha+\beta),$$

即有

$$(\mathscr{A}(\alpha),\mathscr{A}(\alpha))+2(\mathscr{A}(\alpha),\mathscr{A}(\beta))+(\mathscr{A}(\beta),\mathscr{A}(\beta))$$
$$=(\alpha,\alpha)+2(\alpha,\beta)+(\beta,\beta)$$

由$|\mathscr{A}(\alpha)|=|\alpha|$，$|\mathscr{A}(\beta)|=|\beta|$，即有

$$(\mathscr{A}(\alpha),\mathscr{A}(\beta))=(\alpha,\beta),\quad \forall\alpha,\beta\in\mathfrak{L}.$$

这证明了 $\mathscr{A}$ 保持内积不变. 反之,若 $\mathscr{A}$ 保持内积不变,当然它保持长度不变. 证完.

定理 9.1.9 n 维 Euclid 空间 $\mathfrak{R}$ 上线性变换 $\mathscr{A}$ 为正交变换当且仅当在标准正交基 $\{\beta_1,\beta_2,\cdots,\beta_n\}$ 下的方阵表示为 n 阶实正交方阵. 所以在 Euclid 空间 $\mathfrak{R}$ 中取定标准正交基后, 则 $\mathfrak{R}$ 上正交变换和 n 阶实正交方阵间有一个自然的一一对应.

证 记

$$\mathscr{A}(\beta_i)=\sum_{j=1}^{n}a_{ji}\beta_j,\quad i=1,2,\cdots,n.$$

它对应方阵

$$A=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}.$$

由定理 9.1.8, $\mathscr{A}$ 为正交变换当且仅当 $\mathscr{A}$ 保持内积不变,即有

$$(\mathscr{A}(\beta_i),\mathscr{A}(\beta_j))=(\beta_i,\beta_j)=\delta_{ij},\quad 1\leqslant i,j\leqslant n.$$

此即

$$\delta_{ij}=(\mathscr{A}(\beta_i),\mathscr{A}(\beta_j))=\left(\sum_{k=1}^{n}a_{ki}\beta_k,\ \sum_{l=1}^{n}a_{lj}\beta_l\right)$$
$$=\sum_{k,l=1}^{n}a_{ki}a_{lj}(\beta_k,\beta_l)=\sum_{k=1}^{n}a_{ki}a_{kj}.$$

写成矩阵关系,即为

$$A'A=I.$$

所以 $A'=A^{-1}$ 即 $AA'=I$. 这证明了 $\mathscr{A}$ 为正交变换当且仅当在标准正交基下的方阵表示为 n 阶实正交方阵. 证完.

n 维 Euclid 空间 $\mathfrak{R}$ 中任取子集合 $\mathfrak{S}_1,\mathfrak{S}_2$, 则 $(\mathfrak{S}_1,\mathfrak{S}_2)=0$ 表示 $(\alpha,\beta)=0,\ \forall\alpha\in\mathfrak{S}_1,\ \beta\in\mathfrak{S}_2$.

定义 设 $\mathfrak{S}$ 为 n 维 Euclid 空间 $\mathfrak{R}$ 中子集. 记

$$\mathfrak{S}^{\perp}=\{\alpha\in\mathfrak{L}\mid(\alpha,\mathfrak{S})=0\}.$$

则 $\mathfrak{S}^{\perp}$ 称为 $\mathfrak{S}$ 的**正交补**.

定理 9.1.10 n 维 Euclid 空间 $\mathfrak{L}$ 中子集 $\mathfrak{S}$ 的正交补 $\mathfrak{S}^{\perp}$ 为 $\mathfrak{L}$ 的子空间，且

$$\mathfrak{S}\subset(\mathfrak{S}^{\perp})^{\perp}.$$

又

$$\mathfrak{L}=\mathfrak{S}^{\perp}\oplus(\mathfrak{S}^{\perp})^{\perp}$$

为空间直接和. 特别，当 $\mathfrak{S}$ 本身为子空间时，则有

$$(\mathfrak{S}^{\perp})^{\perp}=\mathfrak{S}.$$

所以 $\mathfrak{L}$ 有空间直接和分解

$$\mathfrak{L}=\mathfrak{S}\oplus\mathfrak{S}^{\perp}.$$

证 显然 $\mathfrak{S}^{\perp}$ 为子空间，由 $(\mathfrak{S}^{\perp})^{\perp}$ 之定义可知 $\mathfrak{S}\subset(\mathfrak{S}^{\perp})^{\perp}$. 下面证 $\mathfrak{L}=\mathfrak{S}^{\perp}\oplus(\mathfrak{S}^{\perp})^{\perp}$ 为空间直接和. 任取 $\alpha\in\mathfrak{S}^{\perp}\cap(\mathfrak{S}^{\perp})^{\perp}$，则 $\alpha\in\mathfrak{S}^{\perp}$ 且 $\alpha\in(\mathfrak{S}^{\perp})^{\perp}$，即 $(\alpha,\mathfrak{S}^{\perp})=0$，所以 $(\alpha,\alpha)=0$. 由于内积为定正对称双线性函数，所以 $\alpha=0$. 这证明了 $\mathfrak{S}^{\perp}\cap(\mathfrak{S}^{\perp})^{\perp}=0$. 所以问题化为证 $\mathfrak{S}^{\perp}+(\mathfrak{S}^{\perp})^{\perp}=\mathfrak{L}$，即证 $\dim\mathfrak{S}^{\perp}+\dim(\mathfrak{S}^{\perp})^{\perp}=n$.

在 $\mathfrak{S}^{\perp}$ 中取标准正交基 $\alpha_1,\cdots,\alpha_r$，于是在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$. 由于 $(\alpha_{r+j},\alpha_k)=0$，$j=1,2,\cdots,n-r$，$k=1,2,\cdots,r$，这证明了 $\alpha_{r+j}\in(\mathfrak{S}^{\perp})^{\perp}$，$j=1,\cdots,n-r$. 所以 $\mathfrak{S}^{\perp}+(\mathfrak{S}^{\perp})^{\perp}=\mathfrak{L}$. 由 $\mathfrak{S}^{\perp}\cap(\mathfrak{S}^{\perp})^{\perp}=0$ 便证明了 $\mathfrak{L}=\mathfrak{S}^{\perp}+(\mathfrak{S}^{\perp})^{\perp}$ 为空间直接和. 特别，当 $\mathfrak{S}$ 本身为子空间时，在 $\mathfrak{S}$ 中取定标准正交基 $\beta_1,\cdots,\beta_r$，由定理 9.1.2，在 $\mathfrak{L}$ 中存在标准正交基 $\beta_1,\cdots,\beta_r,\beta_{r+1},\cdots,\beta_n$. 由正交补之定义可知 $\beta_{r+1},\cdots,\beta_n$ 在 $\mathfrak{S}^{\perp}$ 中，所以 $\mathfrak{S}+\mathfrak{S}^{\perp}=\mathfrak{L}$，再 $\mathfrak{S}\cap\mathfrak{S}^{\perp}\subset(\mathfrak{S}^{\perp})^{\perp}\cap\mathfrak{S}^{\perp}=0$，这证明了 $\mathfrak{L}=\mathfrak{S}+\mathfrak{S}^{\perp}$ 为空间直接和. 定理证完.

定义 对 n 维 Euclid 空间 $\mathfrak{L}$ 中子空间 $\mathfrak{S}$，任取 $\alpha\in\mathfrak{L}$，则唯一存在 $\beta\in\mathfrak{S},\gamma\in\mathfrak{S}^{\perp}$，使得 $\alpha=\beta+\gamma$. 于是 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的对应 $\alpha\to\beta$ 为

$\mathfrak{L}$ 上线性变换,称为 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的**正交投影**.

定理9.1.11 设 $\mathscr{A}$ 为 n 维 Euclid 空间 $\mathfrak{L}$ 上线性变换. 记 $\mathscr{A}$ 的象空间 $\mathscr{A}(\mathfrak{L})=\mathfrak{S}$. 则 $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的正交投影当且仅当

$$(\mathscr{A}(\mathfrak{L}),(id-\mathscr{A})(\mathfrak{L}))=0.$$

这时有

$$\mathscr{A}^2=\mathscr{A}.$$

证 今 $\mathfrak{L}=\mathfrak{S}\oplus\mathfrak{S}^{\perp}$. 任取 $\alpha\in\mathfrak{L}$, 记 $\beta\in\mathfrak{S}$, $\gamma\in\mathfrak{S}^{\perp}$, 使得 $\alpha=\beta+\gamma$. 由定义 $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的正交投影当且仅当 $\mathscr{A}(\alpha)=\beta$. 显然 $\mathscr{A}(\alpha)=\beta$ 当且仅当 $(id-\mathscr{A})(\alpha)=\alpha-\beta=\gamma$, 此即

$$(id-\mathscr{A})(\mathfrak{L})=\mathfrak{S}^{\perp}.$$

由 $\mathscr{A}(\mathfrak{L})=\mathfrak{S}$,所以 $\mathscr{A}$ 为正交投影当且仅当 $(\mathscr{A}(\mathfrak{L}),(id-\mathscr{A})(\mathfrak{L}))=0$.

设 $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的正交投影. 于是 $\mathscr{A}^2(\alpha)=\mathscr{A}(\mathscr{A}(\alpha))=\mathscr{A}(\beta)=\beta=\mathscr{A}(\alpha)$, $\forall\alpha\in\mathfrak{L}$, 此即 $\mathscr{A}^2=\mathscr{A}$. 证完.

习题 9.1

1. 设 $\mathscr{A}$ 为 n 维 Euclid 空间 $\mathfrak{L}$ 上保持内积不变的变换. 试证: $\mathscr{A}$ 为正交变换.

2. 试给出一例说明存在实方阵,它的列向量两两正交,但是它的行向量不两两正交.

3 实正交方阵 A 若为准上(下)三角方阵,则对角块都是实正交方阵,且 A 为准对角方阵.

4. 设 A 和 B 为实正交方阵,且 $\det A+\det B=0$. 试证: $A+B$ 为 n 阶奇异方阵.

5. 设 O 为 n 阶实正交方阵, $A=\mathrm{diag}(a_1,\cdots,a_n)$. 试证: 实方阵 OA 的特征根 λ 适合

$$\min_i|a_i|\leqslant|\lambda|\leqslant\max_i|a_i|.$$

6. 试证: 实正交方阵的任一子方阵的特征根之模小于或等于 1.

7. 试证: n 阶实正交方阵必为一些如下形状的 n 阶实正交方阵的乘积

$$\begin{pmatrix} I^{(n-1)} & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} I^{(j-1)} & 0 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta & 0 \\ 0 & 0 & I^{(k-j-1)} & 0 & 0 \\ 0 & -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 0 & I^{(n-k)} \end{pmatrix},$$

其中 $0\leqslant\theta<2\pi, 1\leqslant j<k\leqslant n$.

8. 试证：行列式等于1的三阶实正交方阵 A 必有

$$A=\begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix} \cdot\begin{pmatrix} \cos\xi & \sin\xi & 0 \\ -\sin\xi & \cos\xi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

其中 θ, φ, ξ 称为 Euler 角.

9. 在 n 维 Euclid 空间 Ω 中任取 r 个向量 $\delta_1,\delta_2,\cdots,\delta_r$. 记

$$G(\delta_1,\delta_2,\cdots,\delta_r)=\begin{pmatrix} (\delta_1,\delta_1) & \cdots & (\delta_1,\delta_r) \\ \vdots & & \vdots \\ (\delta_r,\delta_1) & \cdots & (\delta_r,\delta_r) \end{pmatrix}$$

称为 Gram 矩阵. 试证：$\delta_1,\delta_2,\cdots,\delta_r$ 线性无关当且仅当 $\det G(\delta_1,\delta_2,\cdots,\delta_r)\neq 0$. 又任取 Ω 中两组基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 及 $\{\beta_1,\beta_2,\cdots,\beta_n\}$，则

$$G(\alpha_1,\cdots,\alpha_n)=G(\beta_1,\cdots,\beta_n)$$

当且仅当存在正交变换 $\mathscr{A}$，使得 $\mathscr{A}(\alpha_i)=\beta_i, i=1,2,\cdots,n$.

10. 设 n 阶实方阵 A 的秩为 r. 试证：存在实正交方阵 O 及置换方阵 P，使得

$$PAO=\begin{pmatrix} T_1^{(r)} & 0 \\ T_2 & 0 \end{pmatrix},$$

其中 T_1 为 r 阶下三角方阵，对角元素都大于零.

11. 设 λ_0 为 n 阶实正交方阵 O 的特征根，α 为属于特征根 λ_0 的特征向量. 试证：$|\lambda_0|=1$. 设 λ_0 为复数时，$\alpha=\beta+\sqrt{-1}\gamma$ 为复向量，其中 β 和 γ 为实向量. 试证：β 和 γ 长度相等且互相正交.

12. 设 O 为三阶实正交方阵，且 $\det O=1$. 试证：在区间 $[-1,3]$ 中存在实数 λ_0，使得

$$O^3-\lambda_0 O^2+\lambda_0 O-I^{(3)}=0.$$

13. 设K为n阶实斜对称方阵，λ_0为K的非零特征根. 试证：λ_0为纯虚数，记属于λ_0的特征向量$\alpha=\beta+\sqrt{-1}\gamma$，其中$\beta$和$\gamma$为实向量. 试证：$\beta$和$\gamma$长度相等且互相正交.

14. 试求最大个数r，使得n维 Euclid 空间$\mathfrak{L}$中有r个向量$\beta_1,\beta_2,\cdots,\beta_r$，使得$(\beta_i,\beta_j)<0, 1\leqslant i<j\leqslant r$.

15. 设$\mathfrak{L}$为n维 Euclid 空间. 试证：

(i) 在$\mathfrak{L}$中给定非零向量α，则$\mathfrak{L}$上函数$f(\beta)=(\alpha, \beta)$连续，即任取$\varepsilon>0$，则存在$\eta>0$，使得当$\sigma\in\mathfrak{L}$，$|\sigma-\beta|<\eta$时有$|f(\sigma)-f(\beta)|<\varepsilon$.

(ii) 对$\mathfrak{L}$中有限个向量$\beta_1,\cdots,\beta_s$，则存在向量β_0，使得$(\beta_0,\beta_j)\neq 0$，$j=0,1,\cdots,s$;

(iii) 设$\mathscr{A}_1,\mathscr{A}_2,\cdots,\mathscr{A}_s$为$\mathfrak{L}$上有限多个两两不同的线性变换，则存在向量$\beta_0$，使得$\mathscr{A}_1(\beta_0),\mathscr{A}_2(\beta_0),\cdots,\mathscr{A}_s(\beta_0)$也两两不等.

16. 在n维 Euclid 空间中取定有限多个真子空间$\mathfrak{L}_1,\mathfrak{L}_2,\cdots,\mathfrak{L}_r$，试证：存在一组标准正交基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$，使得每个$\alpha_i$都不在$\bigcup\limits_{j=1}^{r}\mathfrak{L}_j$中.

17. 设$\mathscr{A}$为n阶实方阵，且A的特征根不等于$0,-1$. 试证：A及$A+I$都非异，且A实正交当且仅当

$$(A+I)^{-1}+(A'+I)^{-1}=I.$$

§9.2 实方阵在实正交相似下的标准形

定义 n阶实方阵A和B称为**实正交相似的**，如果存在n阶实正交方阵O，使得

$$B=OAO^{-1}=OAO'.$$

显然实正交相似为等价关系.

定理 9.2.1 n阶实方阵A实正交相似于准下(上)三角方阵

$$B=\begin{pmatrix} A_{11}^{(2)} & & & & & \\ \vdots & \ddots & & & & \\ A_{t1}^{(2)} & \cdots & A_{tt}^{(2)} & & & \\ A_{2t+1,1}^{(1,2)} & \cdots & A_{2t+1,t}^{(1,2)} & \lambda_{2t+1} & & \\ \vdots & & \vdots & \vdots & \ddots & \\ A_{n1}^{(1,2)} & \cdots & A_{nt}^{(1,2)} & \lambda_{n,2t+1} & \cdots & \lambda_n \end{pmatrix},$$

其中 $A_{11},\cdots,A_{tt}$ 为二阶实方阵，特征根不是实数. 又 $\lambda_{2t+1},\cdots,\lambda_n$ 为 A 的所有实特征根. 特别，当 A 的特征根都是实数时，A 实正交相似于下(上)三角方阵.

证 若证明了 n 阶实方阵 A 实正交相似于准下三角方阵，则对 A' 用此结果，便证明了 A 实正交相似于准上三角方阵.

为了证明定理，对阶数 n 作归纳法. 当 $n=1$ 时不用再证了. 设对阶数 $\leqslant n-1$ 的实方阵，定理成立. 现在考虑 n 阶实方阵 A. 设 A 有实特征根 λ_n，则属于特征根 λ_n 的特征向量 β_n 仍为实向量. 显然，无妨设 β_n 为单位向量，所以有

$$\beta_n'\beta_n=1,\quad A\beta_n=\lambda_n\beta_n.$$

由定理 9.1.3，存在 n 阶实正交方阵

$$O_1=(\beta_1\beta_2\cdots\beta_n).$$

于是

$$O_1'AO_1=\begin{pmatrix}\beta_1'\\ \vdots\\ \beta_n'\end{pmatrix}A(\beta_1\cdots\beta_n)$$

$$=\begin{pmatrix}\beta_1'A\beta_1 & \cdots & \beta_1'A\beta_n\\ \vdots & & \vdots\\ \beta_n'A\beta_1 & \cdots & \beta_n'A\beta_n\end{pmatrix}.$$

由 $\beta_i'A\beta_n=\lambda_n\beta_i'\beta_n=\delta_{in}\lambda_n$，所以

$$O_1'AO_1=\begin{pmatrix}A_1^{(n-1)} & 0\\ A_2 & \lambda_n\end{pmatrix}.$$

由归纳法假设，存在 $n-1$ 阶实正交方阵 O_2，使得 $O_2'A_1O_2$ 为准下三角方阵，形如定理所给. 而

$$\begin{pmatrix}O_2 & 0\\ 0 & 1\end{pmatrix}'O_1'AO_1\begin{pmatrix}O_2 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}O_2'A_1O_2 & 0\\ A_2O_2 & \lambda_n\end{pmatrix}.$$

这证明了定理. 设 n 阶实正交方阵 A 没有实特征根. 任取一个非实特征根 $\lambda_0+\sqrt{-1}\mu_0$，其中 λ_0,μ_0 都是实数. 自然，$\mu_0\neq0$. 记相

应的特征向量为 $\sigma+\sqrt{-1}\tau$，其中 σ 和 τ 都是实向量. 则由

$$A(\sigma+\sqrt{-1}\tau)=(\lambda_0+\sqrt{-1}\mu_0)(\sigma+\sqrt{-1}\tau),$$

所以有

$$A\sigma=\lambda_0\sigma-\mu_0\tau,\quad A\tau=\mu_0\sigma+\lambda_0\tau.$$

我们有 σ,τ 线性无关. 用反证法. 设若线性相关，则存在非零实向量 β，使得 $\sigma+\sqrt{-1}\tau=a\beta$，其中 a 为非零复常数. 由 $A(\sigma+\sqrt{-1}\tau)=aA\beta$，$(\lambda_0+\sqrt{-1}\mu_0)(\sigma+\sqrt{-1}\tau)=a(\lambda_0+\sqrt{-1}\mu_0)\beta$，所以有 $A\beta=(\lambda_0+\sqrt{-1}\mu_0)\beta$，但是 $A\beta$ 实，$\beta\neq0$，故有 $\mu_0=0$. 这和条件 $\mu_0\neq0$ 矛盾. 今 σ,τ 线性无关，所以在 V_n 中可取基 $\sigma,\tau,\gamma_3,\cdots,\gamma_n$. 对它作 Schmidt 正交化，便得标准正交基 $\beta_1,\beta_2,\cdots,\beta_n$. 使得

$$\beta_{n-1}=b\sigma,$$
$$\beta_n=c\sigma+d\tau,$$

其中 $b>0,d>0$. 因此 $\beta'_j\sigma=\beta'_j\tau=0,\ j=1,2,\cdots,n-2$. 而

$$\begin{aligned}A(\beta_{n-1}\beta_n)&=A(\sigma\ \ \tau)\begin{pmatrix}b&c\\0&d\end{pmatrix}\\&=(\sigma\ \ \tau)\begin{pmatrix}\lambda_0&\mu_0\\-\mu_0&\lambda_0\end{pmatrix}\begin{pmatrix}b&c\\0&d\end{pmatrix}\\&=(\beta_{n-1}\ \ \beta_n)\begin{pmatrix}b&c\\0&d\end{pmatrix}^{-1}\begin{pmatrix}\lambda_0&\mu_0\\-\mu_0&\lambda_0\end{pmatrix}\begin{pmatrix}b&c\\0&d\end{pmatrix}\end{aligned}$$

又当 $j=1,2,\cdots,n-2$，则有

$$\beta'_jA\beta_{n-1}=b\beta'_jA\sigma=b\beta'_j(\lambda_0\sigma-\mu_0\tau)=0,$$
$$\beta'_jA\beta_n=c\beta'_jA\sigma+d\beta'_jA\tau=d\beta'_j(\mu_0\sigma+\lambda_0\tau)=0.$$

记 n 阶实正交方阵，有

$$O_1=(\beta_1\beta_2\cdots\beta_n),$$

则有

$$O_1'AO_1=\begin{pmatrix}\beta_1'A\beta_1 & \cdots & \beta_1'A\beta_n\\ \vdots & & \vdots\\ \beta_n'A\beta_1 & \cdots & \beta_n'A\beta_n\end{pmatrix}$$

$$=\begin{pmatrix}A_1^{(n-2)} & 0\\ A_2 & \begin{pmatrix}\beta_{n-1}'A\beta_{n-1} & \beta_{n-1}'A\beta_n\\ \beta_n'A\beta_{n-1} & \beta_n'A\beta_n\end{pmatrix}\end{pmatrix}$$

$$=\begin{pmatrix}A_1^{(n-2)} & 0\\ A_2 & \begin{pmatrix}\beta_{n-1}'\\ \beta_n'\end{pmatrix}(\beta_{n-1}\ \beta_n)\begin{pmatrix}b & c\\ 0 & d\end{pmatrix}^{-1}\begin{pmatrix}\lambda_0 & \mu_0\\ -\mu_0 & \lambda_0\end{pmatrix}\begin{pmatrix}b & c\\ 0 & d\end{pmatrix}\end{pmatrix}$$

$$=\begin{pmatrix}A_1^{(n-2)} & 0\\ A_2 & \begin{pmatrix}b & c\\ 0 & d\end{pmatrix}^{-1}\begin{pmatrix}\lambda_0 & \mu_0\\ -\mu_0 & \lambda_0\end{pmatrix}\begin{pmatrix}b & c\\ 0 & d\end{pmatrix}\end{pmatrix}.$$

由于A_1的特征根仍为A的特征根．所以A_1仍没有实特征根．今$\begin{pmatrix}b & c\\ 0 & d\end{pmatrix}^{-1}\begin{pmatrix}\lambda_0 & \mu_0\\ -\mu_0 & \lambda_0\end{pmatrix}\begin{pmatrix}b & c\\ 0 & d\end{pmatrix}$的特征根为$\begin{pmatrix}\lambda_0 & \mu_0\\ -\mu_0 & \lambda_0\end{pmatrix}$的特征根，即为$\lambda_0+\sqrt{-1}\mu_0$，$\lambda_0-\sqrt{-1}\mu_0$．对$A_1$用归纳法假设，便证明了定理成立．证完

下面进一步考虑一些特殊的方阵在实正交相似下的标准形．

定理 9.2.2 (Schur 不等式) 记$\lambda_1, \lambda_2, \cdots, \lambda_n$为$n$阶实方阵$A$的$n$个特征根，则有不等式

$$\sum_{j=1}^{n}|\lambda_j|^2\leqslant \mathrm{tr}AA'.$$

且等号成立当且仅当实方阵A适合条件

$$AA'=A'A.$$

这类方阵称为**实规范方阵**.

证 用定理 9.2.1．所以存在n阶实正交方阵O，使得$O'AO=B$．因此

$$\mathrm{tr}AA'=\mathrm{tr}(OBO')(OB'O')=\mathrm{tr}OBB'O^{-1}=\mathrm{tr}BB'.$$

由B为准下三角方阵，所以

$$\mathrm{tr}BB'=\sum_{1\leqslant j\leqslant k\leqslant t}\mathrm{tr}A_{kj}A'_{kj}+\sum_{j=2t+1}^{n}\sum_{k=1}^{t}\mathrm{tr}A_{jk}A'_{jk}$$

$$+\sum_{2t+1\leqslant j<k\leqslant n}\lambda_{kj}^2+\sum_{j=2t+1}^{n}\lambda_j^2$$

$$\geqslant\sum_{j=1}^{t}\mathrm{tr}A_{jj}A'_{jj}+\sum_{j=2t+1}^{n}\lambda_j^2.$$

今A_{jj}的特征根为$\lambda_j,\lambda_{t+j}=\bar{\lambda}_j$. 记$A_{jj}=\begin{pmatrix}a_j & b_j\\ c_j & d_j\end{pmatrix}$，则

$$\mathrm{tr}A_{jj}A'_{jj}=a_j^2+b_j^2+c_j^2+d_j^2\geqslant 2(a_jd_j-b_jc_j)=2\det A_{jj}.$$

所以

$$\mathrm{tr}A_{jj}A'_{jj}\geqslant|\lambda_j|^2+|\lambda_{t+j}|^2,\quad j=1,2,\cdots,t.$$

总之，证明了

$$\mathrm{tr}AA'\geqslant\sum_{j=1}^{n}|\lambda_j|^2.$$

即 Schur 不等式成立.

今等号成立当且仅当

$$\sum_{1\leqslant j<k\leqslant t}\mathrm{tr}A_{kj}A'_{kj}+\sum_{j=2t+1}^{n}\sum_{k=1}^{t}\mathrm{tr}A_{jk}A'_{jk}+\sum_{2t+1\leqslant j<k\leqslant n}\lambda_{kj}^2$$

$$+\sum_{j=1}^{t}(\mathrm{tr}A_{jj}A'_{jj}-|\lambda_j|^2-|\lambda_{t+j}|^2)=0.$$

此即$A_{kj}=0,1\leqslant j<k\leqslant t,A_{jk}=0,1\leqslant k\leqslant t,2t+1\leqslant j\leqslant n,\lambda_{kj}=0,$ $2t+1\leqslant j<k\leqslant n,\mathrm{tr}A_{jj}A'_{jj}=2\det A_{jj},1\leqslant j\leqslant t$. 此即

$$O'AO=\mathrm{diag}(A_{11}^{(2)},\cdots,A_{tt}^{(2)},\lambda_{2t+1},\cdots,\lambda_n).$$

今$A_{jj}=\begin{pmatrix}a_j & b_j\\ c_j & d_j\end{pmatrix}$，$\mathrm{tr}A_{jj}A'_{jj}=2\det A_{jj}$当且仅当

$$(a_j-d_j)^2+(b_j+c_j)^2=0,$$

此即 $d_j=a_j, c_j=-b_j, j=1,2,\cdots,t$. 所以证明了

$$O'AO=\text{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix},\cdots,\begin{pmatrix} a_t & b_t \\ -b_t & a_t \end{pmatrix},\lambda_{2t+1},\cdots,\lambda_n\right).$$

因此

$$AA'=O\text{diag}(a_1^2+b_1^2, a_1^2+b_1^2,\cdots,a_t^2+b_t^2,a_t^2 \\ +b_t^2,\lambda_{2t+1}^2,\cdots,\lambda_n^2)O'.$$

再计算 $A'A$, 便证明了 $AA'=A'A$.

反之, 若 $AA'=A'A$, 于是 $BB'=B'B$, 即有

$$\begin{pmatrix} A_{11} & & & & & & \\ A_{21} & A_{22} & & & & & \\ \vdots & \vdots & \ddots & & & & \\ A_{t1} & A_{t2} & \cdots & A_{tt} & & & \\ A_{2t+1,1} & A_{2t+1,2} & \cdots & A_{2t+1,t} & \lambda_{2t+1} & & \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \\ A_{n1} & A_{n2} & \cdots & A_{nt} & \lambda_{n,2t+1} & \cdots & \lambda_n \end{pmatrix}$$

$$\cdot\begin{pmatrix} A'_{11} & A'_{21} & \cdots & A'_{t1} & A'_{2t+1,1} & \cdots & A'_{n1} \\ & A'_{22} & \cdots & A'_{t2} & A'_{2t+1,2} & \cdots & A'_{n2} \\ & & \ddots & \vdots & \vdots & & \vdots \\ & & & A'_{tt} & A'_{2t+1,t} & \cdots & A'_{nt} \\ & & & & \lambda_{2t+1} & \cdots & \lambda_{n,2t+1} \\ & & & & & \ddots & \vdots \\ & & & & & & \lambda_n \end{pmatrix}$$

$$=\begin{pmatrix} A'_{11} & A'_{21} & \cdots & A'_{t1} & A'_{2t+1,1} & \cdots & A'_{n1} \\ & A'_{22} & \cdots & A'_{t2} & A'_{2t+1,2} & \cdots & A'_{n2} \\ & & \ddots & \vdots & \vdots & & \vdots \\ & & & A'_{tt} & A'_{2t+1,t} & \cdots & A'_{nt} \\ & & & & \lambda_{2t+1} & \cdots & \lambda_{n,2t+1} \\ & & & & & \ddots & \vdots \\ & & & & & & \lambda_n \end{pmatrix}$$

$$
\cdot\begin{pmatrix}
A_{11} & & & & & & \\
A_{21} & A_{22} & & & & & \\
\vdots & \vdots & \ddots & & & & \\
A_{t1} & A_{t2} & \cdots & A_{tt} & & & \\
A_{2t+1,1} & A_{2t+1,2} & \cdots & A_{2t+1,t} & \lambda_{2t+1} & & \\
\vdots & \vdots & & \vdots & \vdots & \ddots & \\
A_{n1} & A_{n2} & \cdots & A_{nt} & \lambda_{n,2t+1} & \cdots & \lambda_n
\end{pmatrix}.
$$

于是有 $A_{11}A'_{11}=A'_{11}A_{11}+\cdots+A'_{t1}A_{+1}+A'_{2t+1,1}A_{2t+1}+\cdots+A'_{n1}A_{n1}$, 双方取迹，便证明了 $A_{21}=0,\cdots,A_{t1}=0,A_{2t+1,1}=0,\cdots,A_{n1}=0$. 同理依次讨论下去，便证明了

$$O'AO=B=\mathrm{diag}(A_{11},\cdots,A_{tt},\lambda_{2t+1},\cdots,\lambda_n),$$

其中 $A_{jj}A'_{jj}=A'_{jj}A_{jj}$. 记 $A_{jj}=\begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix}$, 于是有

$$a_j^2+b_j^2=a_j^2+c_j^2, \quad a_jc_j+b_jd_j=a_jb_j+c_jd_j, \quad c_j^2+d_j^2=b_j^2+d_j^2.$$

于是

$$b_j^2=c_j^2, \quad (a_j-d_j)(c_j-b_j)=0.$$

但是 A_{jj} 的特征根是非实的，所以 $(a_j-d_j)^2+4b_jc_j<0$, 因此 $b_jc_j<0$, 所以 $c_j=-b_j\neq 0$, $d_j=a_j$. 这证明了

$$A_{jj}=\begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}, \quad j=1,2,\cdots,t$$

且

$$\mathrm{tr}AA'=2\sum_{j=1}^{t}(a_j^2+b_j^2)+\sum_{j=2t+1}^{n}\lambda_j^2=\sum_{j=1}^{n}|\lambda_j|^2$$

即 Schur 不等式变为等式. 证完.

上面定理的证明，实际上也给出了

定理 9.2.3 n 阶实规范方阵 A 实正交相似于准对角形

$$B=\mathrm{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix},\cdots,\begin{pmatrix} a_t & b_t \\ -b_t & a_t \end{pmatrix},\lambda_{2t+1},\cdots,\lambda_n\right),$$

其中 $b_1>0,\cdots,b_t>0$. 且 A 的特征根为 $a_1+\sqrt{-1}\,b_1$, $a_1-\sqrt{-1}\,b_1,\cdots,a_t+\sqrt{-1}\,b_t$, $a_t-\sqrt{-1}\,b_t$, λ_{2t+1}, $\cdots$, λ_n. 所以实规范方阵 A 在实正交相似下的全系不变量为 A 的所有特征根.

定理 9.2.4 n 阶实正交方阵 O 实正交相似于准对角形

$$B=\operatorname{diag}\left(\begin{pmatrix}\cos\theta_1 & \sin\theta_1\\ -\sin\theta_1 & \cos\theta_1\end{pmatrix},\cdots,\begin{pmatrix}\cos\theta_t & \sin\theta_t\\ -\sin\theta_t & \cos\theta_t\end{pmatrix},\right.$$
$$\left.I^{(u)},-I^{(v)}\right),$$

其中 $2t+u+v=n,0<\theta_1\leqslant\cdots\leqslant\theta_t<\pi$. 所以实正交方阵的特征根的模为 1，且实正交方阵 O 在实正交相似下的全系不变量为 O 的所有特征根.

证 由 O 实正交，即有 $OO'=O'O=I$，所以它是实规范方阵，由定理 9.2.3，它相似于准对角形

$$B=\operatorname{diag}\left(\begin{pmatrix}a_1 & b_1\\ -b_1 & a_1\end{pmatrix},\cdots,\begin{pmatrix}a_t & b_t\\ -b_t & a_t\end{pmatrix},\lambda_{2t+1},\cdots,\lambda_n\right).$$

由 $\lambda_j^2=1$，所以 $\lambda_j=\pm1$. 无妨取 $\lambda_{2t+1}=\cdots=\lambda_{2t+u}=1$, $\lambda_{2t+u+1}=\cdots=\lambda_n=-1$. 又 $\begin{pmatrix}a_j & b_j\\ -b_j & a_j\end{pmatrix}\begin{pmatrix}a_j & -b_j\\ b_j & a_j\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$，此即

$a_j^2+b_j^2=1$，因此可取 $a_j=\cos\theta_j,b_j=\sin\theta_j$, $0\leqslant\theta_j<\pi$. 证完.

定理 9.2.5 n 阶实对称方阵 S 实正交相似于对角形

$$\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n),$$

其中 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$. 且 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 S 的特征根. 所以实对称方阵的特征根必为实数. 且实对称方阵在实正交相似下的全系不变量为它的所有特征根.

证 由定理 9.2.3，今存在实正交方阵 O，使得

$$O'SO=\operatorname{diag}\left(\begin{pmatrix}a_1 & b_1\\ -b_1 & a_1\end{pmatrix},\cdots,\begin{pmatrix}a_t & b_t\\ -b_t & a_t\end{pmatrix}\lambda_{2t+1},\cdots,\lambda_n\right).$$

由 S 对称可知 $O'SO$ 也对称，所以 $b_1=\cdots=b_t=0$. 这和假设 $b_1>$

$0, \cdots, b_t > 0$ 矛盾. 所以证明了

设 $$O'SO = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n).$$

证完.

定理 9.2.6 n 阶实斜对称方阵 K 实正交相似于准对角形

$$\mathrm{diag}\left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & b_t \\ -b_t & 0 \end{pmatrix}, 0\right),$$

其中 $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_t > 0$, 且 $\sqrt{-1}\, b_1, -\sqrt{-1}\, b_1, \cdots, \sqrt{-1}\, b_t, -\sqrt{-1}\, b_t, 0, \cdots, 0$ 为 K 的特征根. 所以实斜对称方阵的非零特征根必为纯虚数, 因此实斜对称方阵的秩必为偶数. 又实斜对称方阵在实正交相似下的全系不变量为它的所有特征根.

证 由定理 9.2.3, 所以存在实正交方阵 O, 使得

$$O'KO = \mathrm{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_t & b_t \\ -b_t & a_t \end{pmatrix}, \lambda_{2t+1}, \cdots, \lambda_n\right).$$

由 K 斜对称, 所以 $O'KO$ 也是斜对称的. 这证明了

$$a_1 = \cdots = a_t = \lambda_{2t+1} = \cdots = \lambda_n = 0,$$

即有

$$O'KO = \mathrm{diag}\left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & b_t \\ -b_t & 0 \end{pmatrix}, 0, \cdots, 0\right).$$

证完.

引理 9.2.1 设 $\mathfrak{R}$ 为 n 维 Euclid空间, $\mathscr{A}$ 为 $\mathfrak{R}$ 上线性变换, 则它在 $\mathfrak{R}$ 的不同标准正交基下对应的方阵表示互相实正交相似.

证 由于标准正交基间的基变换对应于实正交方阵. 而 $\mathfrak{R}$ 上线性变换在不同基下的方阵表示在基变换对应的方阵下相似. 这证明了引理. 证完.

下面定义共轭变换. 为此, 先给出下面重要定理.

定理 9.2.7 记 (α, β) 为 n 维 Euclid 空间 $\mathfrak{R}$ 上内积. 任取线性函数 f, 则在 $\mathfrak{R}$ 中唯一存在向量 α_f, 使得

$$f(\beta)=(\alpha_f,\beta),\quad \forall\beta\in\mathfrak{L}.$$

且对应 $f\to\alpha_f$ 给出了对偶空间 $\mathfrak{L}^*$ 到 $\mathfrak{L}$ 上的线性同构.

证 在 $\mathfrak{L}$ 中取定标准正交基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$,,任取 $f\in\mathfrak{L}^*$,作

$$\alpha_f=\sum_{j=1}^{n}f(\alpha_j)\alpha_j.$$

则对任意 $\beta=\sum_{k=1}^{n}y_k\alpha_k\in\mathfrak{L}$,有

$$(\alpha_f,\beta)=\sum_{j=1}^{n}f(\alpha_j)y_j=f\left(\sum_{j=1}^{n}y_j\alpha_j\right)=f(\beta).$$

这证明了 α_f 的存在性. 再证唯一性,今若

$$f(\beta)=(\alpha_f,\beta)=(\beta_f,\beta),\forall\beta\in\mathfrak{L}.$$

于是 $(\alpha_f-\beta_f,\beta)=0,\forall\beta\in\mathfrak{L}$, 此即,$\alpha_f-\beta_f\in\mathfrak{L}^{\perp}=0$,即 $\beta_f=\alpha_f$. 这证明了唯一性. 因此对应 $f\to\alpha_f$ 为 $\mathfrak{L}^*$ 到 $\mathfrak{L}$ 内的一一映射. 下面证到上. 今任取 $\alpha\in\mathfrak{L}$,则显然$(\alpha,\beta),\forall\beta\in\mathfrak{L}$ 为 $\mathfrak{L}$ 上线性函数,记作 f,即有

$$f(\beta)=(\alpha,\beta)\qquad\forall\beta\in\mathfrak{L}.$$

下面证 $\alpha_f=\alpha$. 事实上, 已知存在 $\alpha_f\in\mathfrak{L}$, 使得 $f(\beta)=(\alpha_f,\ \beta)$, $\forall\beta\in\mathfrak{L}$. 由唯一性可知 $\alpha_f=\alpha$. 这证明了对应 $f\to\alpha_f$ 为 $\mathfrak{L}^*$ 到 $\mathfrak{L}$ 上的一一对应.

最后证 $f\to\alpha_f$ 为线性同构. 事实上, 任取 $f,\ g\in\mathfrak{L}^*$, $\lambda,\ \mu$ 为数,则

$$(\alpha_{\lambda f+\mu g},\ \beta)=(\lambda f+\mu g)(\beta)=\lambda f(\beta)+\mu g(\beta)=\lambda(\alpha_f,\ \beta)$$
$$+\mu(\alpha_g,\beta)=(\lambda\alpha_f+\mu\alpha_g,\beta),\quad\forall\beta\in\mathfrak{L}.$$

这证明了 $\alpha_{\lambda f+\mu g}=\lambda\alpha_f+\mu\alpha_g$. 即 $f\to\alpha_f$ 为 $\mathfrak{L}^*$ 到 $\mathfrak{L}$ 上的线性同构. 证完.

记$(\alpha,\ \beta)$为 n 维 Euclid 空间 $\mathfrak{L}$ 上的内积. 设 $\mathscr{A}$ 为 $\mathfrak{L}$ 上线性变换. 任给 $\beta\in\mathfrak{L}$,易知 $f(\alpha)=(\mathscr{A}(\alpha),\beta)$, $\forall\alpha\in\mathfrak{L}$ 为线性函数. 由

上面定理可知 $f(\alpha)$ 唯一决定了 Ω 中向量 β_f，使得

$$(\mathscr{A}(\alpha),\beta)=(\alpha,\beta_f).$$

定理 9.2.8 符号同上. 则 $\beta\to\beta_f$ 为 Ω 上线性变换，记作 $\mathscr{A}^*$. 即有

$$(\mathscr{A}(\alpha),\beta)=(\alpha,\mathscr{A}^*(\beta)).$$

$\mathscr{A}^*$ 称为 $\mathscr{A}$ 的**共轭变换**. 在 Ω 中取定标准正交基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$，记线性变换 $\mathscr{A}$ 对应的方阵表示为 A，则共轭变换 $\mathscr{A}^*$ 对应的方阵表示为 A'.

证 先证 $\mathscr{A}^*$ 为线性变换. 任取 $\beta,\gamma\in\Omega$，a,b 为纯量. 于是

$$\begin{aligned}(\mathscr{A}(\alpha),a\beta+b\gamma)&=a(\mathscr{A}(\alpha),\beta)+b(\mathscr{A}(\alpha),\gamma)\\&=a(\alpha,\mathscr{A}^*(\beta)+b(\alpha,\mathscr{A}^*(\gamma))\\&=(\alpha,a\mathscr{A}^*(\beta)+b\mathscr{A}^*(\gamma)).\end{aligned}$$

另一方面，

$$(\mathscr{A}(\alpha),a\beta+b\gamma)=(\alpha,\mathscr{A}^*(a\beta+b\gamma)).$$

因此有

$$(\alpha,\mathscr{A}^*(a\beta+b\gamma))=(\alpha,a\mathscr{A}^*(\beta)+b\mathscr{A}^*(\gamma)),\quad\forall\alpha\in\Omega.$$

这证明了 $\mathscr{A}^*(a\beta+b\gamma)=a\mathscr{A}^*(\beta)+b\mathscr{A}^*(\gamma)$，即 $\mathscr{A}^*$ 为线性变换.

令 $\mathscr{A}(\alpha_i)=\sum_{j=1}^{n}a_{ji}\alpha_j$，$1\leqslant i\leqslant n$，而 $\mathscr{A}$ 在基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 下的方阵表示为

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nn}\end{pmatrix}.$$

于是记 $\mathscr{A}^*(\alpha_i)=\sum_{j=1}^{n}b_{ji}\alpha_j$，$1\leqslant i\leqslant n$，由

$$(\mathscr{A}(\alpha_i),\alpha_j)=(\alpha_i,\mathscr{A}^*(\alpha_j)),\quad i,j=1,2,\cdots,n$$

则有

$$\left(\sum_{k=1}^{n} a_{ki}\alpha_k,\ \alpha_j\right)=\left(\alpha_i,\ \sum_{l=1}^{n} b_{lj}\alpha_l\right).$$

这证明了 $a_{ji}=b_{ij}$, $1\leqslant i,j\leqslant n$. 所以 $\mathscr{A}^*$ 的方阵表示为 $B=A'$. 证完.

定义 记(α,β)为 n 维 Euclid 空间 $\mathfrak{L}$ 的内积. $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 称为**规范的(正交的、对称的、斜对称的)**,如果有 $\mathscr{A}\mathscr{A}^*=\mathscr{A}^*\mathscr{A}$ ($\mathscr{A}\mathscr{A}^*=\mathscr{A}^*\mathscr{A}=id$、$\mathscr{A}^*=\mathscr{A}$、$\mathscr{A}^*=-\mathscr{A}$),即有$(\mathscr{A}(\alpha),\mathscr{A}(\beta))=(\mathscr{A}^*(\alpha),\ \mathscr{A}^*(\beta))$ $((\mathscr{A}(\alpha),\ \mathscr{A}(\beta))=(\alpha,\ \beta)$、$(\mathscr{A}(\alpha),\ \beta)=(\alpha,\mathscr{A}(\beta))$、$(\mathscr{A}(\alpha),\beta)+(\alpha,\mathscr{A}(\beta))=0)$ $\quad\forall\alpha,\beta\in\mathfrak{L}$.

所以定理 9.2.3 到定理 9.2.6 可改写为

定理 9.2.9 记$(\alpha,\ \beta)$为 n 维 Euclid 空间 $\mathfrak{L}$ 的内积, $\mathscr{A}$ 为$\mathfrak{L}$上线性变换. 则有

(1) $\mathscr{A}$为规范变换当且仅当在 $\mathfrak{L}$ 中存在标准正交基$\{\alpha_1,\ \alpha_2,\cdots,\alpha_n\}$,使得

$$\begin{cases}\mathscr{A}(\alpha_1)=a_1\alpha_1-b_1\alpha_2,\\ \mathscr{A}(\alpha_2)=b_1\alpha_1+a_1\alpha_2;\end{cases}\cdots,\quad\begin{cases}\mathscr{A}(\alpha_{2t-1})=a_t\alpha_{2t-1}-b_t\alpha_{2t},\\ \mathscr{A}(\alpha_{2t})=b_t\alpha_{2t-1}+a_t\alpha_{2t},\end{cases}$$

$$\mathscr{A}(\alpha_{2t+1})=\lambda_{2t+1}\alpha_{2t+1},\cdots,\mathscr{A}(\alpha_n)=\lambda_n\alpha_n.$$

其中 $b_1\geqslant\cdots\geqslant b_t>0$.

(2) $\mathscr{A}$为实正交变换当且仅当在 $\mathfrak{L}$ 中存在标准正交基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$,使得

$$\begin{cases}\mathscr{A}(\alpha_1)=(\cos\theta_1)\alpha_1-(\sin\theta_1)\alpha_2,\\ \mathscr{A}(\alpha_2)=(\sin\theta_1)\alpha_1+(\cos\theta_1)\alpha_2;\end{cases}\quad\cdots$$

$$\begin{cases}\mathscr{A}(\alpha_{2t-1})=(\cos\theta_t)\alpha_{2t-1}-(\sin\theta_t)\alpha_{2t}.\\ \mathscr{A}(\alpha_{2t})=(\sin\theta\theta_t)\alpha_{2t-1}+(\cos\theta_t)\alpha_{2t}.\end{cases}$$

$$\mathscr{A}(\alpha_{2t+1})=\alpha_{2t+1},\quad\cdots,\quad\mathscr{A}(\alpha_{2t+u})=\alpha_{2t+u},$$

$$\mathscr{A}(\alpha_{2t+u+1})=-\alpha_{2t+u+1},\cdots,\mathscr{A}(\alpha_n)=-\alpha_n,$$

其中 $$0<\theta_1\leqslant\cdots\leqslant\theta_t<\pi.$$

(3) $\mathscr{A}$为实对称变换当且仅当在Ω中存在标准正交基$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$，使得

$$\mathscr{A}(\alpha_1)=\lambda_1\alpha_1, \cdots, \mathscr{A}(\alpha_n)=\lambda_n\alpha_n.$$

(4) $\mathscr{A}$为实斜对称变换当且仅当在Ω中存在标准正交基$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$，使得

$$\begin{cases}\mathscr{A}(\alpha_1)=-b_1\alpha_2,\\ \mathscr{A}(\alpha_2)=b_1\alpha_1,\end{cases} \cdots, \begin{cases}\mathscr{A}(\alpha_{2t-1})=-b_t\alpha_{2t},\\ \mathscr{A}(\alpha_{2t})=b_t\alpha_{2t-1},\end{cases}$$

$$\mathscr{A}(\alpha_{2t+1})=0, \cdots, \mathscr{A}(\alpha_n)=0,$$

其中
$$b_1\geqslant\cdots\geqslant b_t>0.$$

习题 9.2

1. 给定一组两两可交换的n阶实规范方阵(不一定有限个)．试证：存在一个公共的实正交方阵，将它们同时化为标准形．

2. 设$S_1, \cdots, S_m$为m个n阶实对称方阵，且

$$\sum_{j=1}^{m} S_j^2=0.$$

试证：$S_1=S_2=\cdots=S_m=0$.

3. 试证：实对称方阵的对应不同特征根的特征向量互相正交．因此存在一组标准正交基，全由特征向量构成．

4. 试证：任给n阶实对称方阵S．则存在实对称方阵S_1，使得$S_1^3=S$．它唯一吗？

5. 设A为秩等于r的实对称(实斜对称)方阵，则存在r阶主子式不等于零．

6. 设A和B为n阶实规范方阵．试证：AB为实规范方阵当且仅当BA为实规范方阵．

7. 设O为n阶实正交方阵，特征根不等于-1．试证：$I+O$非异，且

$$K=(I-O)(I+O)^{-1}$$

为实斜对称方阵，且

$$O=(I-K)(I+K)^{-1}.$$

反之，任给实斜对称方阵K．则$I+K$非异，且$O=(I-K)(I+K)^{-1}$实正

交，且 O 的特征根不等于 -1. 因此上面的关系给出了特征根不等于 -1 的实正交方阵和实斜对称方阵间的一一对应，称为 Cayley 变换.

8. 设 A 为实幂等方阵，即 $A^2=A$. 试证：A 只有特征根 $0,1$，且 A 实正交相似于

$$\begin{pmatrix} I^{(r)} & 0 \\ A_1 & 0 \end{pmatrix}.$$

由此证明：存在非异方阵 P，使得

$$P^{-1}AP=\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}$$

9. 实对称方阵 S 为幂等的当且仅当它只有特征根 $0,1$. 且记 $S=(a_{ij})$，则有 $a_{jj}\geqslant 0$，$1\leqslant j\leqslant n$，又 $a_{jj}=0$ 当且仅当 $a_{jk}=a_{kj}=0$，$k=1,2,\cdots,n$.

10. 试证：实规范变换将不变子空间的正交补保持不变.

11. 设 A 和 B 为 n 阶实正交方阵，则 $n-\operatorname{rank}(A+B)$ 为偶数当且仅当 $\det A\det B=1$.

12. 设 A 为 n 阶对合方阵，即 $A^2=I^{(n)}$. 试证：存在 n 阶实正交方阵 O，使得

$$O'AO=\begin{pmatrix} I^{(r)} & 0 \\ A & -I^{(n-r)} \end{pmatrix}.$$

因此存在实非异方阵 P，使得

$$P^{-1}AP=\begin{pmatrix} I^{(r)} & 0 \\ 0 & -I^{(n-r)} \end{pmatrix}.$$

13. 设 A 和 B 为 $n\times m$ 实矩阵，试证有不等式：

$$\operatorname{tr}(AB'-BA')(AB'-BA')'\leqslant \operatorname{tr}AA'\operatorname{tr}BB'.$$

且等式成立当且仅当存在 n 阶实正交方阵 O_1 和 m 阶实正交方阵 O_2，使得

$$O_1AO_2=\begin{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix},\ O_1BO_2\begin{pmatrix} \begin{pmatrix} 0 & -b \\ a & 0 \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix},$$

其中 $a>0$，$b\geqslant 0$.

14. 设 A 为 n 阶实方阵. 试证：A 为对称方阵当且仅当 $A^2=AA'$ 能否给出四种不同的证明？

(i) 用 Cauchy 不等式于 A 的元素上;

(ii) 用迹;

(iii) 用化为非异方阵的情形;

(iv) 用先证明正交相似于上三角方阵.

§9.3 实对称方阵

由定理 9.2.5, 给定 n 阶实对称方阵 S, 则存在实正交方阵

$$O=(\beta_1\cdots\beta_n)$$

使得

$$SO=O\,\mathrm{diag}(\lambda_1,\cdots,\lambda_n),\quad \lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n.$$

所以有

$$S\beta_i=\lambda_i\beta_i,\quad i=1,2,\cdots,n.$$

其中 $\beta_1,\cdots,\beta_n$ 为 V_n 中标准正交基.

定理 9.3.1 (Courant－Fisher 最小最大定理) 设

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$$

为 n 阶实对称方阵 S 的 n 个实特征根, 则

$$\lambda_j=\min_{\forall \xi_1,\cdots,\xi_{j-1}\in V_n}\ \max_{\substack{(x,\xi_k)=0\\ 1\leqslant k<j}}\frac{x'Sx}{x'x},\quad j=1,2,\cdots,n.$$

证 令

$$O'SO=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)=\Lambda,\quad \lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n.$$

记

$$v_k=O\xi_k,\quad 1\leqslant k<j,\ y=Ox$$

于是对 V_n 的标准内积(,)有

$$\mu_j = \min_{\forall \xi_1,\cdots,\xi_{j-1}\in V_n} \max_{\substack{(x,\xi_k)=0\\1\leqslant k<j}} \frac{x'Sx}{x'x} = \min_{\forall v_1,\cdots,v_{j-1}\in V_n}$$

$$\max_{\substack{(y,v_k)=0\\1\leqslant k<j}} \frac{y'\Lambda y}{y'y}.$$

记

$$y=\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix},$$

则

$$\mu_j = \min_{\forall v_1,\cdots,v_{j-1}\in V_n} \max_{\substack{(y,v_k)=0\\1\leqslant k<j}} \frac{\sum_{l=1}^{n}\lambda_k u_k^2}{\sum_{l=1}^{n} u_k^2}$$

记

$$e_j=\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} j, \quad j=1,2,\cdots,n$$

取 $v_1=e_1,\cdots,v_{j-1}=e_{j-1}$，则

$$\mu_j \leqslant \max_{\substack{(y,e_k)=0\\1\leqslant k<j}} \frac{\sum_{l=1}^{n}\lambda_l u_l^2}{\sum_{l=1}^{n} u_l^2} = \max_{u_j,\cdots,u_n\in V_n} \frac{\sum_{k=j}^{n}\lambda_k u_k^2}{\sum_{k=j}^{n} u_k^2} \leqslant \lambda_j.$$

另一方面，任取 $v_1,\cdots,v_{j-1}\in V_n$，若 $y\in V_n$ 有

$$(y,v_k)=0, \quad k=1,2,\cdots,j-1.$$

这是 $j-1$ 个方程，而 $y=\sum_{l=1}^{j} u_l e_l$，即有 j 个独立未知数．所以一定有非零解 $y=y^{(0)}$．于是

$$\mu_j=\min_{\forall v,\cdots,v_{j-1}\in V_n}\ \max_{\substack{(y,v_k)=0\\1\leqslant k<j}}\frac{\sum_{l=1}^{n}\lambda_l u_l^2}{\sum_{l=1}^{n}u_l^2}$$

$$\geqslant \min_{\forall v_1,\cdots,v_{j-1}\in V_n}\frac{\sum_{k=1}^{j}\lambda_k u_k^2}{\sum_{k=1}^{j}u_k^2}$$

$$\geqslant \min_{\forall v_1,\cdots,v_{j-1}\in V_n}\lambda_j=\lambda_j.$$

这证明了 $\mu_j=\lambda_j, j=1,2,\cdots,n$. 证完.

定理 9.3.2 (Courant-Fisher 最大最小定理) 设

$$\lambda_1\geqslant\cdots\geqslant\lambda_n$$

为 n 阶实对称方阵 S 的 n 个特征根，则

$$\lambda_j=\max_{\forall t_1,\cdots,t_{n-j}\in V_n}\ \min_{\substack{(x,t_k)=0\\1\leqslant k\leqslant n-j}}\frac{x'Sx}{x'x}, j=1,2,\cdots,n.$$

证 和定理 9.3.1 的证明相同，我们有

$$\mu_j=\max_{\forall v_1,\cdots,v_{n-j}\in V_n}\ \min_{\substack{(y,v_k)=0\\1\leqslant k\leqslant n-j}}\frac{\sum_{l=1}^{n}\lambda_l u_l^2}{\sum_{l=1}^{n}u_l^2}, i=1,2,\cdots,n.$$

取 $v_1=e_{j+1},\cdots,v_{n-j}=e_n$, $y=\sum_{k=1}^{j}u_k e_k$，则有

$$u_j=\max_{\forall v_1,\cdots,v_{n-j}\in V_n}\ \min_{\substack{(y,v_k)=0\\1\leqslant k\leqslant n-j}}\frac{\sum_{l=1}^{n}\lambda_l u_l^2}{\sum_{l=1}^{n}u_l^2}$$

$$\geqslant \min_{u_1,\cdots,u_j\in V_n} \frac{\sum_{k=1}^{j}\lambda_k u_k^2}{\sum_{k=1}^{j}u_k^2} \geqslant \lambda_j.$$

另一方面，任意取定 $v_1,\cdots,v_{n-j}$，取$y=\sum_{k=j}^{n}u_k e_k$，使得 $(y,v_k)=0$，$k=1,2,\cdots,n-j$. 它是 $n-j+1$ 个未知数，$n-j$ 个方程构成的线性方程组，所以一定有非零解 $y=y^{(0)}$. 于是

$$\mu_j=\max_{\forall v_1,\cdots,v_{n-j}\in V_n} \min_{\substack{(y,v_k)=0\\1\leqslant k\leqslant n-j}} \frac{\sum_{l=1}^{n}\lambda_l u_l}{\sum_{l=1}^{n}u_l^2}$$

$$\leqslant \max_{\forall v_1,\cdots,v_{n-j}\in V_n} \frac{\sum_{k=j}^{n}\lambda_k u_k^2}{\sum_{k=j}^{n}u_k^2} \leqslant \lambda_j.$$

这证明了 $\mu_j=\lambda_j$，$j=1,2,\cdots,n$. 证完.

定理 9.3.3(单调性定理) 设A和B为n阶实对称方阵，且$B\geqslant 0$. 设A的特征根为$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$，$A+B$ 的特征根为$\mu_1\geqslant\mu_2\geqslant\cdots\geqslant\mu_n$，则有

$$\mu_j\geqslant\lambda_j,\qquad j=1,2,\cdots,n.$$

证 今 $x'Bx\geqslant 0$，$\forall x\in V_n$. 所以有

$$\frac{x'(A+B)x}{x'x}=\frac{x'Ax}{x'x}+\frac{x'Bx}{x'x}\geqslant\frac{x'Ax}{x'x},\forall\ 0\neq x\in V_n.$$

任取 $\xi_1,\cdots,\xi_{j-1}\in V_n$，任取 $x\in V_n$，使得

$$(\xi_k,x)=0\qquad k=1,2,\cdots,j-1.$$

则有

$$\max_{\substack{(\xi_k,x)=0\\1\leqslant k<j}} \frac{x'(A+B)x}{x'x}\geqslant\frac{y'Ay}{y'y},\quad\forall\ 0\neq y\in V_n,$$

$$(\xi_{k}, y)=0, \quad 1\leqslant k<j.$$

所以有

$$\max_{\substack{(\xi_k, x)=0\\1\leqslant k<j}}\frac{x'(A+B)x}{x'x}\geqslant\max_{\substack{(\xi_k, x)=0\\1\leqslant k<j}}\frac{x'Ax}{x'x}.$$

同理便证明了

$$\min_{\forall\xi_1,\cdots,\xi_{j-1}\in V_n}\max_{\substack{(\xi_k, x)=0\\1\leqslant k<j}}\frac{x'(A+B)x}{x'x}\geqslant\min_{\forall\xi_1,\cdots,\xi_{j-1}\in V_n}\max_{\substack{(\xi_k, x)=0\\1\leqslant k<j}}\frac{x'Ax}{x'x}.$$

此即 $\mu_j\geqslant\lambda_j, j=1,2,\cdots,n.$ 证完.

引理 9.3.1 设 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$ 为 n 阶实对称方阵 S 的 n 个特征根,则有

$$\lambda_1=\max_{x\neq 0}\frac{x'Sx}{x'x}, \quad \lambda_n=\min_{x\neq 0}\frac{x'Sx}{x'x}.$$

证 今存在 n 阶实正交方阵 O,使得

$$O'SO=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)=\Lambda, \quad \lambda_1\geqslant\cdots\geqslant\lambda_n.$$

于是记 $y=\sum_{j=1}^{n}u_je_j=O'x$,则有

$$u=\max_{x\neq 0}\frac{x'Sx}{x'x}=\max_{y\neq 0}\frac{y'\Lambda y}{y'y}=\max_{y\neq 0}\frac{\sum\lambda_ju_j^2}{\sum u_j^2}\geqslant\frac{\sum\lambda_ju_j^2}{\sum u_j^2}\bigg|_{y=e_1}=\lambda_1$$

而

$$\max_{y\neq 0}\frac{\sum\lambda_ju_j^2}{\sum u_j^2}\leqslant\lambda_1.$$

这证明了 $\mu=\lambda_1$. 又

$$\mu_0=\min_{x\neq 0}\frac{x'Sx}{x'x}=\min_{y\neq 0}\frac{y'\Lambda y}{y'y}=\min_{y\neq 0}\frac{\sum\lambda_ju_j^2}{\sum u_j^2}\leqslant\frac{\sum\lambda_ju_j^2}{\sum u_j^2}\bigg|_{y=e_n}=\lambda_n,$$

而

$$\min_{y\neq 0}\frac{\sum\lambda_ju_j^2}{\sum u_j^2}\geqslant\lambda_n.$$

这证明了 $\mu n=\lambda_n$. 证完.

定理 9.3.4 (严格单调性定理) 设 A 和 B 为 n 阶实对称方阵,且 $B>0$. 设 A 和 $A+B$ 的特征根分别为

$$\lambda_1\geqslant\cdots\geqslant\lambda_n,\quad \mu_1\geqslant\cdots\geqslant\mu_n.$$

则有

$$\lambda_j<\mu_j,\quad j=1,2,\cdots,n.$$

证 设定正对称方阵 B 的特征根为 $\rho_1\geqslant\cdots\geqslant\rho_n>0$, 于是由引理 9.3.1,有

$$\frac{x'(A+B)x}{x'x}=\frac{x'Ax}{x'x}+\frac{x'Bx}{x'x}\geqslant\frac{x'Ax}{x'x}+\rho_n,\ \forall 0\neq x\in V_n.$$

和定理 9.3.3 一样可证 $\mu_j\geqslant\lambda_j+\rho_n>\lambda_j, 1\leqslant j\leqslant n$. **证完.**

定理 9.3.5 (分离性定理) 设

$$S=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}$$

为 n 阶实对称方阵,

$$S_j=\begin{pmatrix}a_{11} & \cdots & a_{1j}\\ \vdots & & \vdots\\ a_{j1} & \cdots & a_{jj}\end{pmatrix},$$

为主子矩阵. 设 S_j 的特征根为

$$\lambda_{1j}\geqslant\lambda_{2j}\geqslant\cdots\geqslant\lambda_{jj},\quad j=1,2,\cdots,n.$$

则有

$$\lambda_{i+1,j+1}\leqslant\lambda_{ij}\leqslant\lambda_{i,j+1},\quad 1\leqslant i,j\leqslant n-1.$$

证 记

$$u_j=\begin{pmatrix}x_1\\ \vdots\\ x_j\end{pmatrix},\ j=1,2,\cdots,n,$$

于是

$$u_{j+1}=\begin{pmatrix}u_j\\x_{j+1}\end{pmatrix}, j=1,2,\cdots,n-1.$$

且

$$\frac{\begin{pmatrix}u_j\\0\end{pmatrix}' S\begin{pmatrix}u_j\\0\end{pmatrix}}{\begin{pmatrix}u_j\\0\end{pmatrix}'\begin{pmatrix}u_j\\0\end{pmatrix}}=\frac{u'_jS_ju_j}{u'_ju_j},\quad j=1,2,\cdots,n.$$

由定理 9.3.1,

$$\lambda_{i,j+1}=\min_{\forall\xi_1,\cdots,\xi_{i-1}\in V_{j+1}}\ \max_{\substack{(u_{j+1},\xi_k)=0\\1\leqslant k<i}}\frac{u'_{j+1}S_{j+1}u_{j+1}}{u'_{j+1}u_{j+1}}$$

而取定 $\xi_1,\cdots,\xi_{i-1}$,则有

$$\max_{\substack{(u_{j+1},\xi_k)=0\\1\leqslant k<i}}\frac{u'_{j+1}S_{j+1}u_{j+1}}{u'_{j+1}u_{j+1}}\geqslant\max_{\substack{\left(\begin{pmatrix}u_j\\0\end{pmatrix},\xi_k\right)=0\\1\leqslant k<i}}\frac{u'_jS_ju_j}{u'_ju_j}.$$

记

$$\xi_k=\begin{pmatrix}\eta_k\\ *\end{pmatrix},\quad \eta_k\in V_j, k=1,2,\cdots,i-1,$$

于是证明了

$$\max_{\substack{(u_{j+1},\xi_k)=0\\1\leqslant k<i}}\frac{u'_{j+1}S_{j+1}u_{j+1}}{u'_{j+1}u_{j+1}}\geqslant\max_{\substack{(u_j,\eta_k)=0\\1\leqslant k<i}}\frac{u'_jS_ju_j}{u'_ju_j}\geqslant\lambda_{ij}$$

后一等式由定理 9.3.1 可知. 今

$$\max_{\substack{(u_{j+1},\xi_k)=0\\1\leqslant k<i}}\frac{u'_{j+1}S_{j+1}u_{j+1}}{u'_{j+1}u_{j+1}}\geqslant\lambda_{ij}.$$

由定理 9.3.1, 由于 $\xi_1,\cdots,\xi_{i-1}$ 可以任取,所以证明了

$$\lambda_{i,j+1}\geqslant\lambda_{ij}.$$

再用定理 9.3.2, 今

$$\lambda_{i+1,j+1}=\max_{\forall \xi_1,\cdots,\xi_{j-i}\in V_{j+1}}\ \min_{\substack{(u_{j+1},\xi_k)=0\\1\leqslant k\leqslant j-i}}\frac{u'_{j+1}S_{j+1}u_{j+1}}{u'_{j+1}u_{j+1}}$$

而取定 $\xi_1,\cdots,\xi_{j-i}$，则由定理 9.3.2，有

$$\min_{\substack{(u_{j+1},\xi_k)=0\\1\leqslant k\leqslant j}}\frac{u_{j+1}S_{j+1}u_{j+1}}{u'_{j+1}u_{j+1}}\leqslant\min_{\substack{\left(\begin{pmatrix}u_j\\0\end{pmatrix},\xi_k\right)=0\\1\leqslant k\leqslant j-i}}\frac{u'_jS_ju_j}{u'_ju_j}$$

$$\leqslant\min_{\substack{(u_j,\eta_k)=0\\1\leqslant k\leqslant j-i}}\frac{u'_jS\ u_j}{u'_ju_j}\leqslant\lambda_{ij}$$

于是

$$\max_{\forall \xi_1,\cdots,\xi_{j-i}\in V_{j+1}}\ \min_{\substack{(u_{j+1},\xi_k)=0\\1\leqslant k\leqslant j-i}}\frac{u'_{j+1}S_{j+1}u_{j+1}}{u'_{j+1}u_{j+1}}\leqslant\lambda_{ij},$$

此即 $\lambda_{i+1,j+1}\leqslant\lambda_{ij}$. 证完.

习题 9.3

设 n 阶实对称方阵 $S=(a_{ij})\geqslant0$，且有

$$\sum_{j=1}^{n}a_{ij}=0,\quad 1\leqslant i\leqslant n,$$

则有

(i) 设 S 的特征根为

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_{n-1}\geqslant\lambda_n=0,$$

则

$$S-\lambda_{n-1}\left(I-\frac{1}{n}ee'\right)\geqslant0,$$

其中

$$e=\sum_{j=1}^{n}e_j=\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$$

(ii) 试证:

$$\lambda_{n+1}\leqslant\left(\frac{n}{n-1}\right)\min(a_{11},\cdots,a_{nn}).$$

§ 9.4 线性不等式

在m维实线性空间V_m中取定基e_1，…，e_m，其中e_i为$m\times 1$矩阵

$$e_i=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},\quad i=1,2,\cdots,m.$$

在n维实线性空间V_n中取定基$f_1,\cdots,f_n$，其中f_j为$n\times 1$矩阵

$$f_j=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},\quad j=1,2,\cdots,n.$$

于是V_m到V_n内的线性映射$\mathscr{A}$和$n\times m$实矩阵A之间有一个自然的一一对应，使得

$$\mathscr{A}(e_i)=\sum_{j=1}^{n}a_{ji}f_j,\quad 1\leqslant i\leqslant m.$$

这里

$$A=\begin{pmatrix}a_{11} & \cdots & a_{1m}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nm}\end{pmatrix}.$$

在这一节中，线性映射$\mathscr{A}$和矩阵A间关系就按上面方式约定。

已知在 V_n 及 V_m 中可以引进如下的标准内积:

$$\left(\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}, \begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\right)_{V_n}=\sum_{j=1}^{n}x_jy_j, \quad \left(\begin{pmatrix}u_1\\ \vdots\\ u_m\end{pmatrix}, \begin{pmatrix}v_1\\ \vdots\\ v_m\end{pmatrix}\right)_{V_m}=\sum_{k=1}^{m}u_kv_k.$$

则 V_n 和 V_m 都是 Euclid 空间. 于是

$$V_n=\mathscr{A}(V_m)\oplus(\mathscr{A}(V_m))^{\perp}$$

为空间直接和.

记

$$P_n=\left\{\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\in V_n \,\middle|\, a_1\geqslant 0,\cdots,a_n\geqslant 0\right\},$$

$$P_n^0=\left\{\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\in P_n \,\middle|\, a_1>0,\cdots,a_n>0\right\}.$$

则 P_n 为 V_n 中第一卦限, P_n^0 为 P_n 的内点集.

定义 m 个未知数 $x_1,\cdots,x_m$ 的关系

$$\sum_{j=1}^{m}a_{ij}x_j-a_j\geqslant 0, \quad 1\leqslant i\leqslant n$$

称为线性不等式.

将线性不等式用矩阵语言来描述. 记

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1m}\\ \vdots&&\vdots\\ a_{n1}&\cdots&a_{nm}\end{pmatrix}, \quad \alpha=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}, \quad x=\begin{pmatrix}x_1\\ \vdots\\ x_m\end{pmatrix}.$$

于是线性不等式可改写为

$$AX-\alpha\in P_n.$$

线性不等式的问题是求解:

定义 如果存在实数 $x_1^{(0)},\cdots,x_m^{(0)}$, 使得

$$\sum_{j=1}^{m} a_{ij}x_j^{(0)} - a_i \geqslant 0, \quad i=1,2,\cdots,n$$

则

$$x^{(0)} = \begin{pmatrix} x_1^{(0)} \\ \vdots \\ x_m^{(0)} \end{pmatrix}$$

称为线性不等式 $Ax-\alpha\in P_n$ 的一组解.

定理 9.4.1 记$\mathscr{A}$为 V_m 到 V_n 内的线性映射，它的矩阵表示. A 为 $n\times m$ 实矩阵，α 为 $n\times 1$ 实矩阵. x为m个独立未知数构成的 $m\times 1$ 实矩阵. 则线性不等式

$$Ax-\alpha\in P_n$$

有解当且仅当

$$(P_n\cap(\mathscr{A}(V_m))^{\perp},\alpha)_{V_n}\leqslant 0.$$

证 设线性不等式 $Ax-\alpha\in P_n$ 有解 $x=x^{(0)}$，则 $Ax^{(0)}-\alpha\in P_n$. 任取 $\beta\in P_n\cap(\mathscr{A}(V_m))^{\perp}$，则由 $\beta\in P_n$ 有$(\beta, Ax^{(0)}-\alpha)_{V_n}\geqslant 0$. 因此

$$(\beta,\alpha)_{V_n}\leqslant(\beta, Ax^{(0)})_{V_n}.$$

由 $\beta\in(\mathscr{A}(V_m))^{\perp}$，所以$(\beta, Ax^{(0)})_{V_n}=0$. 这证明了$(\beta,\alpha)_{V_n}\leqslant 0$. 由 β 任取，所以$(P_n\cap(\mathscr{A}(V_m))^{\perp},\alpha)_{V_n}\leqslant 0$.

反之，用反证法，即若不存在 $x^{(0)}\in V_m$，使得 $Ax^{(0)}-\alpha\in P_n$，则必存在 $\beta\in P_n\cap(\mathscr{A}(V_m))^{\perp}$，使得$(\beta,\alpha)_{V_n}>0$.

由条件，不存在 $x^{(0)}\in V_m$，使得 $Ax^{(0)}-\alpha\in P_n$，即

$$\{\mathscr{A}(V_m)-\alpha\}\cap P_n=\varnothing.$$

在 V_n 的子空间$(\mathscr{A}(V_m))^{\perp}$中取基 β_1, $\beta_2,\cdots$, β_r, 在 V_r 中取子集合

$$Q=\{\xi=((\delta,\beta_1)_{V_n},\cdots,(\delta,\beta_r)_{V_n})'\in V_r \mid \forall\delta\in P_n\}$$

于是 Q 是以原点为顶点的闭凸锥.（定义：V_r 中子集$\mathfrak{S}$称为**凸集**，如

果 $\xi,\eta\in\mathfrak{S}$，则 $t\xi+(1-t)\eta\in\mathfrak{S}$，$0\leqslant t\leqslant 1$；$\mathfrak{S}$ 称为锥，如果 $\xi\in\mathfrak{S}$，则 $t\xi\in\mathfrak{S}$，对一切非负实数 t 成立；$\mathfrak{S}$称为闭集，如果 $\mathfrak{S}$ 中收敛序列的极限点仍在 $\mathfrak{S}$ 中）. 证明如下：

今 $\xi,\eta\in Q$，则存在 $\delta,\sigma\in P_n$，使得

$$\xi=((\delta,\beta_1)_{V_n},\cdots,(\delta,\beta_r)_{V_n})',\quad \eta((\sigma,\beta_1)_{V_n},\cdots,(\sigma,\beta_r)_{V_n})'.$$

于是
$$t\xi+(1-t)\eta=((t\delta+(1-t)\sigma,\beta_1)_{V_n},\cdots,(t\delta+(1-t)\sigma,\beta_r)_{V_n})',$$

当 $0\leqslant t\leqslant 1$，则由 $\sigma,\delta\in P_n$ 可知 $t\delta+(1-t)\sigma\in P_n$，所以 $t\xi+(1-t)\eta\in Q$. 这证明了 Q 为凸集. 又当 $t\geqslant 0$，则 $t\xi=((t\delta,\beta_1)_{V_n},\cdots,(t\delta,\beta_r)_{V_n})'$. 由 $\delta\in P_n$ 可知 $t\delta\in P_n$，所以 $t\xi\in Q$. 这证明了 Q 为锥. 最后由 P_n 为闭集，所以 Q 也是闭集，至此证明了 Q 为闭凸锥.

而条件 $\{\mathscr{A}(V_m)-\alpha\}\cap P_n=\varnothing$ 推出

$$\zeta=((-\alpha,\beta_1)_{V_n},\cdots,(-\alpha,\beta_r)_{V_n})'\notin Q.$$

事实上，若 $\zeta\in Q$，即 $((-\alpha,\beta_1)_{V_n},\cdots,(-\alpha,\beta_r)_{V_n})'\in Q$，则存在 $\delta\in P_n$，使得 $((-\alpha,\beta_1)_{V_n},\cdots,(-\alpha,\beta_r)_{V_n})'=((\delta,\beta_1)_{V_n},\cdots,(\delta,\beta_r)_{V_n})'$，即有

$$(\delta+\alpha,\beta_j)_{V_n}=0,\quad j=1,2,\cdots,r.$$

由于 $\beta_1,\cdots,\beta_r$ 为 $(\mathscr{A}(V_m))^{\perp}$ 之基，所以证明了 $\delta+\alpha\in\mathscr{A}(V_m)$. 因此 $\delta\in\{\mathscr{A}(V_m)-\alpha\}\cap P_n$，这和 $\{\mathscr{A}(V_m)-\alpha\}\cap P_n=\varnothing$ 矛盾. 所以证明了 $\zeta\notin Q$.

现在在 V_r 中有闭凸锥 Q 及不在 Q 中之点 ζ. 由图形可以看出：在 V_r 中存在**超平面**

$$\pi:\quad \sum_{i=1}^{r}c_jz_j=c>0$$

使得 π 和 Q 及 ζ 都不交，且将它们隔开. 由于 V_r 中超平面将 V_r 分成两部分，一部分使得 $\sum_{j=1}^{r}c_jz_j-c>0$，另一部分使得 $\sum_{j=1}^{r}c_jz_j-$

$c<0$. 又由于原点在闭凸锥 Q 中,而当 $z_1=\cdots=z_r=0$ 时,

$\sum_{j=1}^{r} c_j z_j - c<0$,这证明了 Q 中点

$$\xi=((\delta,\beta_1)_{V_n},\cdots,(\delta,\beta_r)_{V_n})'$$

代入,仍有

$$\sum_{j=1}^{r} c_j(\delta,\beta_j)_{V_n}-c<0, \forall\delta\in P_n.$$

又 ξ 在超平面 π 的另一侧,所以有

$$\sum_{j=1}^{r} c_j(-\alpha,\beta_j)_{V_n}-c>0.$$

记

$$-\sum_{j=1}^{r} c_j\beta_j=\beta,$$

则有

$$(\delta,\beta)_{V_n}+c>0,\quad (\alpha,\beta)_{V_n}>c>0,\quad \forall\delta\in P_n.$$

这证明了 $(\alpha,\beta)_{V_n}>0$. 余下证 $\beta\in P_n\cap(\mathscr{A}(V_m))^{\perp}$. 由于 $\beta_1,\cdots,\beta_r$ 为 $(\mathscr{A}(V_m))^{\perp}$ 之基,所以 $\beta\in(\mathscr{A}(V_m))^{\perp}$. 最后证 $\beta\in P_n$. 事实上,记

$$\beta=\begin{pmatrix} b_1\\ \vdots\\ b_n\end{pmatrix},\quad \delta=\begin{pmatrix}0\\ \vdots\\ 0\\ \lambda\\ 0\\ \vdots\\ 0\end{pmatrix} j,\quad \lambda\geqslant 0,\ 1\leqslant j\leqslant n.$$

则 $\delta\in P_n$,而

$$0<c+(\delta,\beta)_{V_n}=c+\lambda b_j,\quad \lambda\geqslant 0,\quad 1\leqslant j\leqslant n.$$

若某个 $b_{j_0}<0$, 取 $\lambda>\frac{c}{-b_{j_0}}$, 则 $c+\lambda b_{j_0}<c-c=0$, 这和条件 $c+\lambda b_{j_0}>0$ 矛盾. 所以证明了 β 之坐标都非负, 即 $\beta\in P_n$. 至此证明了存在 $\beta\in P_n\cap(\mathscr{A}(V_m))^{\perp}$, 使得 $(\alpha,\beta)_{V_n}>0$. 证完.

定理 9.4.2 设 A 为 $n\times m$ 实矩阵, α 为 $n\times 1$ 实矩阵, x 为 m 个独立未知数构成的 $m\times 1$ 实矩阵. 则线性不等式

$$Ax-\alpha\in P_n^0$$

有解当且仅当

$$(\{P_n\cap(\mathscr{A}(V_m))^{\perp}\}-\{0\},\alpha)_{V_n}<0.$$

特别, 线性不等式

$$Ax\in P_n^0$$

有解当且仅当

$$P_n\cap(\mathscr{A}(V_m))^{\perp}=\{0\}.$$

证 若存在 $x^{(0)}\in V_m$, 使得 $Ax^{(0)}-\alpha\in P_n^0$. 任取 $\beta\in P_n-\{0\}$, 则有

$$(Ax^{(0)}-\alpha,\beta)_{V_n}=(Ax^{(0)},\beta)_{V_n}-(\alpha,\beta)_{V_n}>0.$$

当 $\beta\in\{P_n\cap(\mathscr{A}(V_m))^{\perp}\}-\{0\}$, 则有 $\beta\in(\mathscr{A}(V_m))^{\perp}$, 于是 $(Ax^{(0)},\beta)_{V_n}=0$. 因此证明了

$$(\alpha,\beta)_{V_n}<0,\forall\beta\in\{P_n-(\mathscr{A}(V_m))^{\perp}\}-\{0\}.$$

反之, 若 $Ax-\beta\in P_n^0$ 无解, 记

$$e=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}=\sum_{j=1}^{n}e_j,$$

则 $Ax-\alpha-\frac{1}{p}e\in P_n$ 无解, 其中 $p=1,2,\cdots$. 用定理 9.4.1 的证明可知: 存在 $\beta^{(p)}\in P_n\cap(\mathscr{A}(V_m))^{\perp}$, 使得

$$\left(\beta^{(p)},\alpha+\frac{1}{p}e\right)_{V_n}>0,\quad p=1,2,\cdots,$$

显然，无妨设 $\beta^{(p)}$ 为单位向量，因此它们都在单位球面

$$S=\{\beta\in V_n \mid \beta'\beta=1\}$$

上，因此 S 上序列 $\{\beta^{(p)}\}_{p=1}^{\infty}$ 中存在收敛子序列 $\{\beta^{(p_j)}\}_{j=0}^{\infty}$，使得它的极限为 β_0. β_0 仍在单位球面上，即为单位向量，又

$$\lim_{j\to\infty}\beta^{(p_j)}=\beta_0.$$

另一方面，由 $\beta^{(p_j)}\in P_n\cap(\mathscr{A}(V_m))^{\perp}$ 及 P_n, $(\mathscr{A}(V_m))^{\perp}$ 都是 V_n 中闭子集，所以极限 $\beta_0\in P_n\cap(\mathscr{A}(V_m))^{\perp}$. 而

$$\left(\beta^{(p_j)},\quad \alpha+\frac{1}{p_j}e\right)_{V_n}>0, j=1,2,\cdots.$$

当 $j\to\infty$，便有 $(\beta_0,\alpha)\geqslant 0$. 这证明了在 $\{P_n\cap((\mathscr{A}V_m))^{\perp}\}-\{0\}$ 中存在向量 β_0，使得 $(\beta_0,\alpha)_{V_n}\geqslant 0$. 所以当 $(\{P_n-(\mathscr{A}(V_m))^{\perp}\}-\{0\},\alpha)_{V_n}<0$，则线性不等式 $Ax-\alpha\in P_n^0$ 有解. 证完.

定理 9.4.3 设 A 为 $n\times m$ 实矩阵，x 为由 m 个独立自变量构成的 $m\times 1$ 实矩阵. 则线性不等式

$$Ax\in P_n$$

在 $P_m-\{0\}$ 中无解当且仅当存在 $u\in P_n^0$，使得

$$-A'u\in P_m^0.$$

证 设 $Ax\in P_n$ 在 $P_m-\{0\}$ 中有解 v，且存在 $u\in P_n^0$，使得 $-A'u\in P_m^0$. 于是

$$v'(A'u)=v'A'u=u'(Av)\geqslant 0.$$

但是 $-A'u\in P_m^0$, $v\in P_m-\{0\}$，所以 $v'(A'u)=(-v)'(-A'u)<0$. 这导出矛盾. 因此当存在 $u\in P_n^0$，使得 $-A'u\in P_m^0$，则 $Ax\in P_n$ 在 $P_m-\{0\}$ 中必无解.

反之，若不存在 $u\in P_n^0$，使得 $-A'u\in P_m^0$. 因此

$$\begin{pmatrix}-A'u\\ u\end{pmatrix}=\begin{pmatrix}-A'\\ I\end{pmatrix}u\in P_{n+m}^0$$

无解. 记$\mathscr{A}$为V_n到V_{n+m}内之线性映射,使得对V_n及V_{n+m}之标准基,$\mathscr{A}$所对应之矩阵表示为$(n+m)\times n$矩阵

$$\begin{pmatrix} -A' \\ I^{(n)} \end{pmatrix},$$

其中$I^{(n)}$为n阶单位方阵. 于是$\mathscr{A}$的共轭变换$\mathscr{A}^*$所对应之矩阵表示为$n\times(n+m)$矩阵

$$\begin{pmatrix} -A' \\ I \end{pmatrix}' = (-A\ I).$$

而任取$\sigma\in V_{n+m}$, $\tau\in V_n$, 则有

$$(\sigma, \mathscr{A}(\tau))_{V_{n+m}}(\mathscr{A}^*(\sigma), \tau)_{V_n}.$$

由定理 9.4.2, $\begin{pmatrix} -A' \\ I \end{pmatrix}u\in P^0_{n+m}$无解当且仅当

$$P_{n+m}\cap(\mathscr{A}(V_n))^{\perp}\neq 0.$$

所以存在$0\neq v\in P_{n+m}\cap(\mathscr{A}(V_n))^{\perp}$. 记

$$v=\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\neq 0$$

其中$w_1\in P_m$, $w_2\in P_n$, 则$\left(\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \mathscr{A}(V_n)\right)_{V_{n+m}}=0$, 即$\left(\mathscr{A}^*\begin{pmatrix} w_1 \\ w_2 \end{pmatrix},\right.$ $\left.V_n\right)_{V_n}=0$, 所以

$$\mathscr{A}^*\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}=0,$$

即

$$(-A\ I)\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}=0,$$

所以$w_2=Aw_1$. 由于$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\neq 0$, 所以$w_1\neq 0$. 因为若$w_1=0$, 则$w_2=Aw_1=0$, 这导出矛盾. 由$w_1\neq 0$可知$w_1\in P_m-\{0\}$, $w_2=Aw_1\in P_n$. 这证明了线性不等式$Ax\in P_n$在$P_n-\{0\}$中有解, 所以这导出矛盾. 定理证完.

第十章　二次型的分类

§10.1　实对称方阵在实相合下的标准形

定义　n 阶实方阵 A 和 B 称为**实相合**的，如果存在 n 阶可逆实方阵 P，使得

$$B=P'AP.$$

引理 10.1.1　实相合为等价关系

证　显然 $A=I'AI$，其中 I 为单位方阵，即 A 和 A 实相合．这证明了反身性．又若 $B=P'AP$，则有 $A=(P^{-1})'B(P^{-1})$，所以 A 和 B 实相合，则 B 和 A 实相合．这证明了对称性．最后若 $B=P'AP, C=Q'BQ$，其中 P,Q 都非异，则 $C=Q'BQ=Q'(P'AP)Q=(PQ)'A(PQ)$．即若 A 和 B 实相合，B 和 C 实相合，则 A 和 C 实相合．这证明了传递性．证完．

显然实对称方阵（实斜对称方阵）在实相合下仍变为实对称方阵（实斜对称方阵）．

在这一章给出实对称方阵及实斜对称方阵在实相合下的标准形及全系不变量，从而给出了二次型的分类．

定理10.1.1　任一实对称方阵 S 在实相合下的标准形为

$$\operatorname{diag}(I^{(p)},-I^{(r-p)},0),$$

其中 $r=\operatorname{rank}(S)$．数值

$$\delta(S)=p-(r--p)=2p-r$$

称为 S 的**符号差**，它是标准形的迹，也是 S 的正特征根个数减负特征根的个数．

证　下面给出两种证明．第二种方法实质上是二次型的**配**

方法.

（一）由定理 9.2.5,存在实正交方阵 O,使得

$$O'SO=\mathrm{diag}(\lambda_1,\cdots,\lambda_r,0,\cdots,0)$$

其中 $r=\mathrm{rank}(S)$, $\lambda_1\geqslant\cdots\geqslant\lambda_p>0>\lambda_{p+1}\geqslant\cdots\geqslant\lambda_r$. 取

$$P_0=\mathrm{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_p},\sqrt{-\lambda_{p+1}},\cdots,\sqrt{-\lambda_r},1,\cdots,1)^{-1},$$

于是 n 阶非异方阵 $P=OP_0$,有

$$P'SP=\mathrm{diag}(I^{(p)},-I^{(r-p)},0).$$

证完.

（二）当 $S=0$,则不必讨论. 当 $S\neq0$,记

$$S=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix},\qquad a_{ij}=a_{ji}.$$

设有一个元素 $a_{jj}\neq0$,作置换方阵 P_{1j},则有

$$P'_{1j}SP_{1j}=\begin{pmatrix}a_{jj}&*\\ *&*\end{pmatrix}.$$

所以在实相合意义下无妨设 $a_{11}\neq0$. 设$a_{11},a_{22},\cdots a_{nn}$都等于零,则存在一个 $a_{ij}\neq0$,其中 $i<j$,作置换方阵 P_{1j},P_{2i},则有

$$P_{2i}'P_{1j}'SP_{1j}P_{2i}=\begin{pmatrix}0&a_{ij}&*\\ a_{ij}&0&*\\ *&*&*\end{pmatrix},$$

所以在实相合意义下无防设 $a_{12}\neq0$. 再

$$\begin{pmatrix}\begin{pmatrix}1&-1\\1&1\end{pmatrix}&0\\ 0&I\end{pmatrix}'\begin{pmatrix}\begin{pmatrix}0&a_{12}\\a_{12}&0\end{pmatrix}&*\\ *&*\end{pmatrix}\begin{pmatrix}\begin{pmatrix}1&-1\\1&1\end{pmatrix}&0\\ 0&I\end{pmatrix}$$

$$=\begin{pmatrix}\begin{pmatrix}2a_{12}&0\\0&-2a_{12}\end{pmatrix}&*\\ *&*\end{pmatrix}.$$

所以在实相合意义下总可无妨设 $a_{11}\neq0$. 这时记

$$S=\begin{pmatrix} a_{11} & \alpha' \\ \alpha & S_1 \end{pmatrix}$$

其中 α 为 $(n-1)\times 1$ 实矩阵, S_1 为 $n-1$ 阶实对称方阵, 作

$$\begin{pmatrix} 1 & -a_{11}^{-1}\alpha' \\ 0 & I^{(n-1)} \end{pmatrix}'\begin{pmatrix} a_{11} & \alpha' \\ \alpha & S_1 \end{pmatrix}\begin{pmatrix} 1 & -a_{11}^{-1}\alpha' \\ 0 & I^{(n-1)} \end{pmatrix}=\begin{pmatrix} a_{11} & 0 \\ 0 & S_1-a_{11}^{-1}\alpha\alpha' \end{pmatrix}$$

再作

$$\begin{pmatrix} \frac{1}{\sqrt{|a_{11}|}} & 0 \\ 0 & I^{(n-1)} \end{pmatrix}'\begin{pmatrix} a_{11} & 0 \\ 0 & S_1-a_{11}^{-1}\alpha\alpha' \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{|a_{11}|}} & 0 \\ 0 & I^{(n-1)} \end{pmatrix}$$

$$=\begin{pmatrix} \pm 1 & 0 \\ 0 & S_1-a_{11}^{-1}\alpha\alpha' \end{pmatrix}.$$

由于对 $S_1-a_{11}^{-1}\alpha\alpha'$ 作实相合可扩充为对 $\begin{pmatrix} \pm 1 & 0 \\ 0 & S_1-a_{11}^{-1}\alpha\alpha' \end{pmatrix}$ 作实相合. 由归纳法便证明了定理. 证完.

引理10.1.2(Witt 定理) 给定 n 阶实对称方阵 S 和 m 阶实对称方阵 S_1 和 S_2, 设 S, S_1, S_2 都非异, 且 $n+m$ 阶实对称方阵

$$\begin{pmatrix} S & 0 \\ 0 & S_1 \end{pmatrix}, \begin{pmatrix} S & 0 \\ 0 & S_2 \end{pmatrix}$$

互相实相合, 则 S_1 和 S_2 实相合.

证 今存在 n 阶非异实方阵 P_1, 使得

$$P_1' S P_1=\begin{pmatrix} I^{(p)} & 0 \\ 0 & -I^{(n-p)} \end{pmatrix}=\begin{pmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_n \end{pmatrix},$$

于是

$$\begin{pmatrix} P_1 & 0 \\ 0 & I^{(m)} \end{pmatrix}'\begin{pmatrix} S & 0 \\ 0 & S_i \end{pmatrix}\begin{pmatrix} P_1 & 0 \\ 0 & I^{(m)} \end{pmatrix}=\begin{pmatrix} P_1' SP & 0 \\ 0 & S_i \end{pmatrix}, \quad i=1,2.$$

记

$$\tilde{S}=\begin{pmatrix} \mathrm{diag}(\delta_2,\cdots,\delta_n) & 0 \\ 0 & S_i \end{pmatrix}, \quad i=1,2.$$

所以上面证明了

$$\begin{pmatrix} \delta_1 & 0 \\ 0 & \tilde{S}_1 \end{pmatrix}, \begin{pmatrix} \delta_1 & 0 \\ 0 & \tilde{S}_2 \end{pmatrix}$$

互相实相合，此即

$$\begin{pmatrix} -\delta_1 & 0 \\ 0 & -\tilde{S}_1 \end{pmatrix}, \begin{pmatrix} -\delta_1 & 0 \\ 0 & -\tilde{S}_2 \end{pmatrix}$$

互相实相合，显然 $\tilde{S}_1$ 和 $\tilde{S}_2$ 实相合当且仅当 $-\tilde{S}_1$ 和 $-\tilde{S}_2$ 实相合。因此我们无妨设 $\delta_1=1$，然后证明 $\tilde{S}_1$ 和 $\tilde{S}_2$实相合。

今若存在 $n+m$ 阶非异实方阵

$$P=\begin{pmatrix} a & \beta' \\ \alpha & P_0 \end{pmatrix},$$

其中 a 为实数，α, β 为$(n+m-1)\times 1$ 实矩阵，P_0 为$n+m-1$ 阶实方阵，使得

$$P'\begin{pmatrix} 1 & 0 \\ 0 & \tilde{S}_1 \end{pmatrix} P=\begin{pmatrix} 1 & 0 \\ 0 & \tilde{S}_2 \end{pmatrix},$$

即有

$$\begin{pmatrix} 1 & 0 \\ 0 & \tilde{S}_2 \end{pmatrix}=\begin{pmatrix} a & \alpha' \\ \beta & P_0' \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \tilde{S}_1 \end{pmatrix}\begin{pmatrix} a & \beta' \\ \alpha & P_0 \end{pmatrix}$$

$$=\begin{pmatrix} a^2+\alpha'\tilde{S}_1\alpha & a\beta'+\alpha'\tilde{S}_1P_0 \\ a\beta+P_0'\tilde{S}_1\alpha & \beta\beta'+P_0'\tilde{S}_1P_0 \end{pmatrix}$$

所以

$$a^2+\alpha'\tilde{S}_1\alpha=1,$$

$$a\beta+P_0'\tilde{S}_1\alpha=0,$$

$$\beta\beta'+P_0'\tilde{S}_1P_0=\tilde{S}_2.$$

显然无妨设 $a\neq 1$，因为若 $a=1$，考虑 $-P$，则有

$$(-P)'\begin{pmatrix} 1 & 0 \\ 0 & \tilde{S}_1 \end{pmatrix}(-P)=\begin{pmatrix} 1 & 0 \\ 0 & \tilde{S}_2 \end{pmatrix},$$

而 $-P$ 之第一行第一列元素为 -1.

取 $n+m-1$ 阶实方阵

$$Q=P_0+(1-a)^{-1}\alpha\beta',$$

则有

$$\begin{aligned}Q'\tilde{S}_1Q&=(P_0'+(1-a)^{-1}\beta\alpha')\tilde{S}_1(P_0+(1-a)^{-1}\alpha\beta')\\&=P_0'\tilde{S}_1P_0+(1-a)^{-1}\beta\alpha'\tilde{S}_1P_0+(1-a)^{-1}P_0'\tilde{S}_1\alpha\beta'\\&\quad+(1-a)^{-2}\beta\alpha'\tilde{S}_1\alpha\beta'\\&=(\tilde{S}_2-\beta\beta')+(1-a)^{-1}\beta(-a\beta')\\&\quad+(1-a)^{-1}(-a\beta)\beta'+(1-a)^{-2}\beta(1-a^2)\beta'\\&=\tilde{S}_2.\end{aligned}$$

由假设 S, S_1, S_2 非异,所以 $\delta_1\det\tilde{S}\neq0$, $\delta_1\det\tilde{S}_2\neq0$. 这证明了 $\det\tilde{S}_1\det\tilde{S}_2\neq0$. 由 $Q'\tilde{S}_1Q=\tilde{S}_2$,双方取行列式,便证明了 $\det Q\neq0$, 因此 $\tilde{S}_1$ 和 $\tilde{S}_2$ 实相合.

对 S 的阶数 n 作归纳法,便证明了 S_1 和 S_2 实相合. 引理证完.

定理 10.1.2(惯性定理) 两个 n 阶方阵

$$A=\begin{pmatrix}I^{(p)}&0&0\\0&-I^{(r-p)}&0\\0&0&0\end{pmatrix},\quad B=\begin{pmatrix}I^{(q)}&0&0\\0&-I^{(s-q)}&0\\0&0&0\end{pmatrix}$$

实相合当且仅当 $r=s$, $p=q$, 所以 n 阶实对称方阵在实相合下的全系不变量为它的秩和符号差.

证 由于 A 和 B 实相合,所以存在 n 阶非异实方阵 P, 使得 $B=P'AP$,所以 $\mathrm{rank}(A)=\mathrm{rank}(B)$,即有 $r=s$.

将 P 按前 r 行,列分块为

$$P=\begin{pmatrix}P_1^{(r)}&P_2\\P_3&P_4\end{pmatrix},$$

于是

$$\begin{pmatrix} \begin{pmatrix} I^{(q)} & 0 \\ 0 & -I^{(r-q)} \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} P_1' & P_3' \\ P_2' & P_4' \end{pmatrix}$$

$$\begin{pmatrix} \begin{pmatrix} I^{(p)} & 0 \\ 0 & -I^{(r-p)} \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix}.$$

这证明了

$$\begin{pmatrix} I^{(q)} & 0 \\ 0 & -I^{(r-q)} \end{pmatrix} = P_1' \begin{pmatrix} I^{(p)} & 0 \\ 0 & -I^{(r-p)} \end{pmatrix} P_1.$$

双方取行列式，便证明了 $\det P_1 \neq 0$，即证明了

$$\begin{pmatrix} I^{(q)} & 0 \\ 0 & -I^{(r-q)} \end{pmatrix}, \begin{pmatrix} I^{(p)} & 0 \\ 0 & -I^{(r-p)} \end{pmatrix}$$

实相合．由 Witt 定理，当 $p>q$ 时证明了

$$\begin{pmatrix} I^{(p-q)} & 0 \\ 0 & -I^{(r-p)} \end{pmatrix}, -I^{(r-q)}$$

互相实相合．所以存在 $r-q$ 阶非异实方阵 Q，使得

$$\begin{pmatrix} I^{(p-q)} & 0 \\ 0 & -I^{(r-p)} \end{pmatrix} = Q'(-I^{(r-q)})Q = -Q'Q.$$

记 $Q=(q_{ij})$，比较第一行第一列元素，便有 $1=-\sum_{i=1}^{r-q} q_{i1}^2$，这导出矛盾，所以 $p \leqslant q$．当 $p<q$ 时同理可导出矛盾．至此证明了 $p=q$．定理证完．

下面的矩阵计算技巧，实际上是配方法的推广，非常有用．

设 S 为 n 阶实对称方阵．将它按前 s 行，列分块为

$$S = \begin{pmatrix} S_{11}^{(s)} & S_{12} \\ S_{12}' & S_{22} \end{pmatrix}.$$

假设 $\det S_{11} \neq 0$，即 S_{11} 为 s 阶非异实对称方阵，则有

$$\begin{pmatrix} I & -S_{11}^{-1}S_{12} \\ 0 & I \end{pmatrix}'\begin{pmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{pmatrix}\begin{pmatrix} I & -S_{11}^{-1}S_{12} \\ 0 & I \end{pmatrix}$$

$$=\begin{pmatrix} I & 0 \\ -S_{12}'S_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{pmatrix}\begin{pmatrix} I & -S_{11}^{-1}S_{12} \\ 0 & I \end{pmatrix}$$

$$=\begin{pmatrix} S_{11} & S_{12} \\ 0 & S_{22}-S_{12}'S_{11}^{-1}S_{12} \end{pmatrix}\begin{pmatrix} I & -S_{11}^{-1}S_{12} \\ 0 & I \end{pmatrix}$$

$$=\begin{pmatrix} S_{11} & 0 \\ 0 & S_{22}-S_{12}'S_{11}^{-1}S_{12} \end{pmatrix}.$$

所以将 n 阶实对称方阵打了两个洞，且证明了 S 实相合于准对角形.

例 设 S 为 n 阶实对称方阵，α 为 $n\times 1$ 实矩阵. 假设 S 和 $S-\alpha\alpha'$ 都非异. 试证:它们的符号差有关系

$$\delta(S)=\begin{cases}\delta(S-\alpha\alpha'), & \text{当 } \alpha'S^{-1}\alpha<1; \\ \delta(S-\alpha\alpha')+2, & \text{当 } \alpha'S^{-1}\alpha>1.\end{cases}$$

证 考虑 $n+1$ 阶实对称方阵

$$\begin{pmatrix} 1 & \alpha' \\ \alpha & S \end{pmatrix}.$$

由

$$\begin{pmatrix} 1 & 0 \\ -\alpha & I \end{pmatrix}\begin{pmatrix} 1 & \alpha' \\ \alpha & S \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -\alpha & I \end{pmatrix}'=\begin{pmatrix} 1 & 0 \\ 0 & S-\alpha\alpha' \end{pmatrix},$$

$$\begin{pmatrix} 1 & -\alpha'S^{-1} \\ 0 & I \end{pmatrix}\begin{pmatrix} 1 & \alpha' \\ \alpha & S \end{pmatrix}\begin{pmatrix} 1 & -\alpha'S^{-1} \\ 0 & I \end{pmatrix}'=\begin{pmatrix} 1-\alpha'S^{-1}\alpha & 0 \\ 0 & S \end{pmatrix}.$$

于是符号差有关系

$$\delta\left(\begin{pmatrix} 1 & 0 \\ 0 & S-\alpha\alpha' \end{pmatrix}\right)=\delta\left(\begin{pmatrix} 1-\alpha'S^{-1}\alpha & 0 \\ 0 & S \end{pmatrix}\right).$$

由于符号差为标准形中 1 之个数减去 -1 之个数，即等于标准形的对角元素相加. 因此证明了：当 $\alpha'S^{-1}\alpha\neq 1$,则

$$1+\delta(S-\alpha\alpha')=\frac{1-\alpha'S^{-1}\alpha}{|1-\alpha'S^{-1}\alpha|}+\delta(S)$$

当 $\alpha'S^{-1}\alpha<1$，则有

$$1+\delta(S-\alpha\alpha')=1+\delta(S);$$

当 $\alpha'S^{-1}\alpha>1$，则有

$$1+\delta(S-\alpha\alpha')=-1+\delta(S).$$

最后，当 $\alpha'S^{-1}\alpha=1$，则有 $\det\begin{pmatrix}1 & \alpha' \\ \alpha & S\end{pmatrix}=0$，即 $\det(S-\alpha\alpha')=0$，这和假设矛盾. 证完.

运用上面理论，下面给出实二次型的结果. 由 §8.1，所以实二次型为

$$\varphi(x)=\sum_{i,j=1}^{n} a_{ij}x_i x_j=x'Sx$$

其中 $S=(a_{ij})$ 为 n 阶实对称方阵. 作坐标变换

$$x_i=\sum_{j=1}^{n} p_{ij}y_j,\quad i=1,2,\cdots n,$$

其中 n 阶实方阵

$$P=(p_{ij})$$

非异. 于是 $x=Py$，因此

$$\varphi(x)=x'Sx=(Py)'S(Py)=y'(P'SP)y.$$

即作坐标变换 $x=Py$，则二次型关于坐标 x 及 y 分别对应实对称方阵 S 及 $P'SP$，它们互相实相合，所以定理 10.1.1 给出了

定理 10.1.3 任给实二次型

$$\varphi(x)=\sum_{i,j=1}^{n} a_{ij}x_i x_j,\quad a_{ij}=a_{ji}$$

则存在 n 个独立未知数 $x_1,\cdots,x_n$ 的非异线性变换

$$x_i=\sum_{j=1}^{n}p_{ij}y_j,\quad i=1,2,\cdots,n$$

使得 $\varphi(x)$ 变为 n 个独立未知数 $y_1,\cdots,y_n$ 的二次型

$$\psi(y)=\sum_{j=1}^{p}y_j^2-\sum_{j=p+1}^{r}y_j^2.$$

将定理 10.1.1 的第二个证明用坐标写出，便得到所谓配方法.

习题 10.1

1. 设 S 为 n 阶实对称方阵

$$S=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix}$$

设 j 阶主子方阵

$$S_j=\begin{pmatrix}a_{11}&\cdots&a_{1j}\\ \vdots&&\vdots\\ a_{j1}&\cdots&a_{jj}\end{pmatrix},\quad j=1,2,\cdots,n$$

都非异，则实二次型

$$\varphi(x)=x'Sx$$

可化为二次型

$$\psi(y)=(\det S_1)y_1^2+\left(\frac{\det S_2}{\det S_1}\right)y_2^2+\cdots+\left(\frac{\det S_n}{\det S_{n-1}}\right)y_n^2.$$

2. 试给出实二次型能分解为两个实线性函数之乘积的必要且充分条件.

3. 设实二次型 $\varphi(x)=x'Sx$ 有性质

$$\det S\neq 0,\ \ \varphi(x_1,\cdots,x_k,0,\cdots,0)\equiv 0.$$

试证：φ 可经过非异线性变换变为二次型

$$y_1y_{k+1}+\cdots+y_ky_{2k}+\psi(y_{2k+1},\cdots,y_n),$$

其中 ψ 对应非异实对称方阵，其符号差之绝对值不超过 $n-2k$.

4. 设 $n\geqslant 3$，且 A 和 B 为 n 阶实对称方阵．设 $n\times 1$ 实对称方阵 α 有 $\alpha'A\alpha=0,\alpha'B\alpha=0$，则必有 $\alpha=0$．试证下面定理：存在 n 阶实非异方阵 P，

使得$P'AP, P'BP$同时化为对角形.

定理的证明可分为下面八步:

1) 所给条件在实相合下不变;

2) 任取 j, k, 其中 $1\leqslant j<k\leqslant n$. 记 $A=(a_{ij})$, $B=(b_{ij})$. 则对任意 $(k-j+1)\times 1$ 实矩阵 β, 由

$$\beta'\begin{pmatrix} a_{jj} & \cdots & a_{jk} \\ \vdots & & \vdots \\ a_{kj} & \cdots & a_{kk} \end{pmatrix}\beta=0, \quad \beta'\begin{pmatrix} b_{jj} & \cdots & b_{jk} \\ \vdots & & \vdots \\ b_{kj} & \cdots & b_{kk} \end{pmatrix}\beta=0$$

可推出 $\beta=0$;

3) 记 $r=\mathrm{rank}(A)$, 则存在 n 阶实非异方阵 P_1, 使得

$$P_1'AP_1=\begin{pmatrix} A_1^{(r)} & 0 \\ 0 & 0 \end{pmatrix}, \quad P_1'BP_1=\begin{pmatrix} B_1^{(r)} & 0 \\ 0 & \pm I \end{pmatrix};$$

4) 在实相合下无妨取

$$A=\begin{pmatrix} A_0^{(p)} & 0 & 0 \\ 0 & I^{(q)} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B=\begin{pmatrix} B_0^{(p)} & 0 & \\ 0 & 0^{(q)} & 0 \\ 0 & 0 & I \end{pmatrix},$$

其中 $\det A_0 \det B_0\neq 0$;

5) 在 $p=2$ 时可直接证明 A 和 B 同时实相合于标准形. 在 $p>2$ 时试将问题化为当 A 和 B 都是非异的情形;

6) 设 A 和 B 都是非异方阵. 试证: $\det(A-\lambda B)$ 为 λ 的实多项式, 它的根也都是实数;

7) 取定 $\det(A-\lambda B)$ 的一个实根 λ_0, 试利用 λ_0 来证明定理成立.

8) 试举反例说明: 任给两实对称方阵 A 和 B, 一般说来, 它们不能同时相合于对角形.

§10.2 实定正对称方阵和实方阵的极分解

显然实对称方阵的定正(半定正、定负、半定负)的性质在实相合下不改变.

定理 10.2.1 设 S 为 n 阶实对称方阵, 则下面性质相互等价:

(1) $S>0$;

(2) S 的特征根都是正实数;

(3) 存在非异实方阵 Q 使得 $S=Q'Q$;

(4) S 的 n 个顺序主子式 $S\begin{pmatrix}12 & \cdots & j\\ 12 & \cdots & j\end{pmatrix}>0, j=1,2,\cdots,n$;

(5) S 的一切主子式都是正实数.

证 (1)⇒(2). 设 $S>0$, 即对一切 $x\in V_n, x\neq 0$, 有 $x'Sx>0$. 另一方面, 由定理 9.2.5, 存在 n 阶实正交方阵 O, 使得 $O'SO=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$, 其中 $\lambda_1\geqslant\cdots\geqslant\lambda_n$. 取

$$e_i=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},\quad e_i\in V_n,\quad i=1,2,\cdots,n.$$

令 $x=Oe_i$, 则 $e_i'O'SOe_i=\lambda_i>0, i=1,2,\cdots,n$. 这证明了 S 的 n 个特征根都大于零. 所以由(1)推出(2).

(2)⇒(3). 假设 n 阶实对称方阵 S 的 n 个特征根 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 都是正实数, 则存在 n 阶实正交方阵 O, 使得 $O'SO=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$. 令

$$Q=\operatorname{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})O',\quad \det Q\neq 0.$$

于是 $S=Q'Q$. 所以由(2)推出(3).

(3)⇒(4). 假设 n 阶实对称方阵 $S=Q'Q, \det Q\neq 0$. 由定理 4.2.1(Binet-Cauchy 公式), 所以有

$$S\binom{12\cdots j}{12\cdots j}=\sum_{1\leqslant i_1<\cdots<i_j\leqslant n}Q'\binom{12\cdots j}{i_1i_2\cdots i_j}Q\binom{i_1i_2\cdots i_j}{12\cdots j}$$

$$=\sum_{1\leqslant i_1<\cdots<i_j\leqslant n}Q\binom{i_1i_2\cdots i_j}{12\cdots j}^2.$$

今 Q 非异, 所以前 j 列线性无关, 因此存在 j 阶子式 $Q\binom{i_1i_2\cdots i_j}{12\cdots j}\neq 0$.

这证明了 $S\binom{12\cdots j}{12\cdots j}>0, j=1,2,\cdots,n$. 所以由(3)推出(4).

(4)⇒(1). 设 n 阶实对称方阵 S 的 n 个顺序主子式大于零. 下面用归纳法来证明 $S>0$. 当 $n=1$ 时，由行列式大于零可知它定正. 设 $n-1$ 阶实对称方阵的 $n-1$ 个顺序主子式都大于零，则它定正. 现在设 n 阶实对称方阵 S 的 n 个顺序主子式大于零. 记

$$S=\begin{pmatrix} S_1{}^{(n-1)} & \alpha \\ \alpha' & a \end{pmatrix},$$

则 S_1 的 $n-1$ 个顺序主子式都大于零，所以 $S_1>0$. 又 S 的第 n 个顺序主子式为 $\det S>0$. 我们来证明 $S>0$. 今

$$\begin{pmatrix} I^{(n-1)} & 0 \\ -\alpha' S_1{}^{-1} & 1 \end{pmatrix}\begin{pmatrix} S_1 & \alpha \\ \alpha' & a \end{pmatrix}\begin{pmatrix} I^{(n-1)} & 0 \\ -\alpha' S_1{}^{-1} & 1 \end{pmatrix}'=\begin{pmatrix} S_1 & 0 \\ 0 & a-\alpha' S_1{}^{-1}\alpha \end{pmatrix},$$

双方取行列式，便有 $\det S=(a-\alpha' S_1{}^{-1}\alpha)\det S_1>0$. 由 $S_1>0$ 可知 $\det S_1>0$,所以证明了 $a-\alpha' S_1{}^{-1}\alpha>0$.

现在任取 $x\in V_n$，记 $x=\begin{pmatrix} y \\ x_n \end{pmatrix}$, 其中 $y\in V_{n-1}$，x_n 为实数. 于是

$$x'\begin{pmatrix} S_1 & 0 \\ 0 & a-\alpha' S_1{}^{-1}\alpha \end{pmatrix}x=y'S\cdot y+(a-\alpha' S_1{}^{-1}\alpha)x_n{}^2.$$

设$x_n\neq 0$，则 $y'S_1y\geqslant 0$，$(a-\alpha' S_1{}^{-1}\alpha)x_n{}^2>0$; 设 $x_n=0$，$y\neq 0$，则 $y'S_1y>0$, $(a-\alpha' S_1{}^{-1}\alpha)x_n{}^2\geqslant 0$. 总之,只要 $x\neq 0$,便有 $x'\mathrm{diag}(S_1, a-\alpha' S_1{}^{-1}\alpha)x>0$,所以 $\mathrm{diag}(S_1, a-\alpha' S_1{}^{-1}\alpha)>0$. 而

$$\begin{pmatrix} S_1 & \alpha \\ \alpha' & a \end{pmatrix}=\begin{pmatrix} I^{(n-1)} & 0 \\ \alpha' S_1{}^{-1} & 1 \end{pmatrix}\begin{pmatrix} S_1 & 0 \\ 0 & a-\alpha' S_1{}^{-1}\alpha \end{pmatrix}\begin{pmatrix} I^{(n-1)} & 0 \\ \alpha' S_1{}^{-1} & 1 \end{pmatrix}'>0.$$

这证明了 $S>0$. 因此由(4)推出(1).

最后证(1)⇒(5), (5)⇒(4). 由上面已证(1)和(4)等价,这样定理也就证明了. 显然(5)可推出(4). 下面证(1)可推出(5). 今 $S>0$，任取 $1\leqslant i_1<\cdots<i_j\leqslant n$, 取 S 的主子式 $S\binom{i_1 i_2\cdots i_j}{i_1 i_2\cdots i_j}$. 记置换方阵 $P=P_{1i_1}P_{2i_2}\cdots P_{ji_j}$, 则有 $P'SP>0$，而 $P'SP$ 的第 j 个顺序主

子式即为$S\begin{pmatrix} i_1 & i_2 & \cdots & i_j \\ i_1 & i_2 & \cdots & i_j \end{pmatrix}$. 由$P'SP>0$可推出$P'SP$的第$j$个顺序主子式大于零,这证明了$S\begin{pmatrix} i_1 & i_2 & \cdots & i_j \\ i_1 & i_2 & \cdots & i_j \end{pmatrix}>0$. 至此证明了当$S$定正,则$S$的任意主子式大于零,即(5)成立. 定理证完.

定理10.2.2 设S为n阶实对称方阵,则下面性质互相等价:

(1) $S\geqslant 0$;

(2) S的特征根都是非负实数;

(3) 存在n阶实方阵Q使得$S=Q'Q$;

(4) S的一切主子式都是非负实数.

证 和定理10.2.1 一样可证(1)⇒(2)⇒(3)⇒(4). 余下证由(4)可推出(1). 今假设S的一切主子式都非负. 记S的特征多项式

$$f(\lambda)=\det(\lambda I-S)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_n,$$

其中$(-1)^j a_j$为S的所有j阶主子式之和. 所以

$$(-1)^j a_j\geqslant 0, \quad j=1,2,\cdots,n.$$

我们来证S的特征多项式$f(\lambda)$的根都是非负实数. 事实上,$f(\lambda)$若有负根$-x_0$,其中$x_0>0$. 则有

$$(-1)^n x_0^n+(-1)^{n-1}a_1x_0^{n-1}+\cdots+(-1)a_{n-1}x_0+a_n=0,$$

此即

$$x_0^n+(-a_1)x_0^{n-1}+\cdots+(-1)^{n-1}a_{n-1}x_0+(-1)^n a_n=0.$$

由于$-a_1\geqslant 0,\cdots,(-1)^{n-1}a_{n-1}\geqslant 0,(-1)^n a_n\geqslant 0$, $x_0>0$, 这导出矛盾. 所以证明了S的特征根都是非负实数. 因此存在实正交方阵O使得

$$S=O'\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)O,$$

其中$\lambda_1\geqslant\cdots\geqslant\lambda_n\geqslant 0$.

由于任取$x\in V_n$, $y=Ox$, 其中

$$y=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

则

$$x'Sx=(Ox)'\mathrm{diag}(\lambda_1,\cdots,\lambda_n)(Ox)=\sum_{j=1}^{n}\lambda_j y_j^2\geqslant 0,$$

这证明了 $S\geqslant 0$，定理证完.

定理10.2.3 设 S 为 n 阶半定正实对称方阵，记 $\mathrm{rank}(S)=r$. 则存在 r 个线性无关的 $n\times 1$ 实矩阵 $\beta_1,\beta_2,\cdots,\beta_r$，使得

$$S=\sum_{j=1}^{r}\beta_j\beta_j'.$$

证 今 $S\geqslant 0$，由定理 10.2.2，存在实正交方阵 O，使得

$$\begin{aligned}S&=O'\mathrm{diag}(\lambda_1,\cdots,\lambda_n)O\\&=\sum_{j=1}^{n}\lambda_j O'\mathrm{diag}(0,\cdots,0,\overset{(j)}{1},0,\cdots,0)O\\&=\sum_{j=1}^{n}\lambda_j O'e_je_j'O,\end{aligned}$$

其中

$$e_j=\begin{pmatrix}0\\\vdots\\0\\1\\0\\\vdots\\0\end{pmatrix}j,\quad e_j\in V_n,\ j=1,2,\cdots,n.$$

今 $\mathrm{rank}(S)=r$，所以 $\lambda_1\geqslant\cdots\geqslant\lambda_r>0,\ \lambda_{r+1}=\cdots=\lambda_n=0$，记

$$\beta_j=\sqrt{\lambda_j}O'e_j,\quad j=1,2,\cdots,r$$

显然 $\beta_1,\beta_2,\cdots,\beta_r$ 线性无关，且

$$S=\sum_{j=1}^{r}\beta_j\beta_j'$$

证完.

例 给定 n 阶定正实对称方阵 A 和 B，记作

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}.$$

则 n 阶实对称方阵

$$C=\begin{pmatrix} a_{11}b_{11} & \cdots & a_{1n}b_{1n} \\ \vdots & & \vdots \\ a_{n1}b_{n1} & \cdots & a_{nn}b_{nn} \end{pmatrix}$$

也定正.

证 今 $a_{ij}=a_{ji}, b_{ij}=b_{ji}$，于是

$$a_{ij}b_{ij}=a_{ji}b_{ji},$$

即 C 为实对称方阵. 由 $A>0, B>0$，所以任取 $x\in V_n, x\neq 0$，则有

$$x'Ax>0, \quad x'Bx>0.$$

下面给出三种方法来证明 n 阶实对称方阵 C 定正.

(一) 今 $B>0$，所以存在 n 阶实非异方阵 Q 使得 $B=Q'Q$. 记

$$Q=\begin{pmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & & \vdots \\ q_{n1} & \cdots & q_{nn} \end{pmatrix}$$

于是

$$b_{ij}=\sum_{k=1}^{n} q_{ki}q_{kj}, \quad 1\leqslant i, j\leqslant n.$$

而任取 $x\in V_n$，则

$$x'Cx=\sum_{i,j=1}^{n} a_{ij}b_{ij}x_ix_j=\sum_{i,j,k=1}^{n} a_{ij}q_{ki}q_{kj}x_ix_j$$

$$=\sum_{k=1}^{n}\sum_{i,j=1}^{n} a_{ij}(q_{ki}x_i)(q_{kj}x_j).$$

由 $A>0$ 可知 $x'Cx\geqslant 0$，且等号成立当且仅当

$$q_{ki}x_i=0,\quad 1\leqslant i,k\leqslant n.$$

记 $x=\mathrm{diag}(x_1,\cdots,x_n)$，因此 $Qx=0$. 但是 $\det Q\neq 0$，所以 $x=0$ 这证明了 $C>0$.

（二）今 $A>0,B>0$. 所以存在 n 阶实非异方阵 P,Q，使得

$$A=P'P,\quad B=Q'Q.$$

记 $A=(a_{ij}),B=(b_{ij}),P=(p_{ij}),Q=(q_{ij})$，于是

$$a_{ij}=\sum_{k=1}^{n}p_{ki}p_{kj},\quad b_{ij}=\sum_{k=1}^{n}q_{ki}q_{kj},\quad 1\leqslant i,j\leqslant n$$

因此任取 $x\in V_n$，则

$$x'Cx=\sum_{i,j=1}^{n}a_{ij}b_{ij}x_ix_j=\sum_{i,j=1}^{n}\sum_{k,l=1}^{n}p_{ki}p_{kj}q_{pi}q_{pj}x_ix_j$$

$$=\sum_{k+p=1}^{n}\left(\sum_{i=1}^{n}p_{ki}q_{pi}x_i\right)^2\geqslant 0.$$

等号成立当且仅当 $\sum_{i=1}^{n}p_{ki}q_{li}x_i=0,k,l=1,2,\cdots,n$. 此即

$$P\begin{pmatrix}x_1 & & \\ & \ddots & \\ & & x_n\end{pmatrix}Q'=0.$$

由于 $\det P\neq 0,\det Q\neq 0$，所以 $x_1=\cdots=x_n=0$，即 $x=0$. 这证明了 $C>0$.

（三）今 $A>0,B>0$，记 $A=(a_{ij})$，$B=(b_{ij})$. 于是存在 n 阶实正交方阵 O 使得

$$A=O'\mathrm{diag}(\lambda_1,\cdots,\lambda_n)O,$$

其中

$$\lambda_1\geqslant\cdots\geqslant\lambda_n>0.$$

记

$$B_1=\text{diag}(x_1,\cdots,x_n)'B\ \text{diag}(x_1,\cdots,x_n)=(d_{ij}),$$

于是 $d_{ij}=b_{ij}x_ix_j$, $1\leqslant i,j\leqslant n$, 且 $B_1\geqslant 0$, 又

$$x'Cx=\sum_{i,j=1}^{n}a_{ij}b_{ij}x_ix_j=\sum_{i,j=1}^{n}a_{ij}d_{ij}=\text{tr}AB_1.$$

而

$$\text{tr}AB_1=\text{tr}O'\text{diag}(\lambda_1,\cdots,\lambda_n)OB_1=\text{tr}\ \text{diag}(\lambda_1,\cdots,\lambda_n)OB_1O'.$$

记

$$OB_1O'=B_2=(e_{ij})\geqslant 0,$$

则 $e_{ii}\geqslant 0, 1\leqslant i\leqslant n$. 因此

$$\text{tr}AB_1=\sum_{j=1}^{n}\lambda_je_{ii}\geqslant 0.$$

而等号成立当且仅当 $e_{11}=\cdots=e_{nn}=0$, 即 $\text{tr}B_2=0$. 今

$$0=\text{tr}B_2=\text{tr}OB_1O'=\text{tr}B_1=\text{tr}B\text{diag}(x_1{}^2,\cdots,x_n{}^2)=\sum_{j=1}^{n}b_{jj}x_j{}^2.$$

而 $B>0$, 所以 $b_{11}>0,\cdots,b_{nn}>0$. 这证明了 $x_1=\cdots=x_n=0$, 即 $x=0$. 所以证明了 $C>0$. 证完.

下面给出极分解.

定理 10.2.4 设 S 为 n 阶半定正实对称方阵, 则唯一存在 n 阶半定正实对称方阵 S_1, 使得

$$S_1{}^2=S.$$

且任一 n 阶实方阵 A, 若有 $AS=SA$, 则必有 $AS_1=S_1A$. 再 S 定正当且仅当 S_1 定正.

证 今存在 n 阶实正交方阵 O, 使得

$$O'SO=\text{diag}(\lambda_1I^{(n_1)},\cdots,\lambda_sI^{(n_s)})=\Lambda,$$

其中

$$\lambda_1>\cdots>\lambda_s\geqslant 0.$$

记

$$\Lambda^{\frac{1}{2}}=\mathrm{diag}(\sqrt{\lambda_1}I^{(n_1)},\cdots,\sqrt{\lambda_s}I^{(n_s)}),$$

$$S_1=O\Lambda^{\frac{1}{2}}O'.$$

则 S_1 为 n 阶实半定正对称方阵，且

$$S_1^2=(O\Lambda^{\frac{1}{2}}O')^2=O\Lambda O'=S.$$

下面证唯一性. 今若另外存在一个 n 阶实半定正对称方阵 S_2，使得 $S_2^2=S$. 于是存在 n 阶实正交方阵 O_1，使得

$$O_1'S_2O_1=\mathrm{diag}(\mu_1I^{(m_1)},\cdots,\mu_tI^{(m_t)}),$$

其中

$$\mu_1>\cdots>\mu_t\geqslant 0.$$

于是

$$\begin{aligned}S=S_2^2&=(O_1\mathrm{diag}(\mu_1I^{(m_1)},\cdots,\mu_tI^{(m_t)})O_1')^2\\&=O_1\mathrm{diag}(\mu_1^2I^{(m_1)},\cdots,\mu_t^2I^{(m_t)})O_1'\\&=O\mathrm{diag}(\lambda_1I^{(n_1)},\cdots,\lambda_s^{(1)}I^{(n_s)})O'.\end{aligned}$$

由于 $\lambda_1,\cdots,\lambda_s$ 为 S 的所有不同特征根，重数分别为 $n_1,\cdots,n_s$；又 $\mu_1^2,\cdots,\mu_t^2$ 为 S 的所有不同特征根，重数分别为 $m_1,\cdots,m_t$. 这证明了 $t=s, m_j=n_j, 1\leqslant j\leqslant s$, 又 $\lambda_j=\mu_j^2, 1\leqslant j\leqslant n$. 这证明了

$$S_2=O_1\Lambda^{\frac{1}{2}}O_1'.$$

由

$$S=O_1\Lambda O_1'=O\Lambda O',$$

所以 $O_1'O\Lambda=\Lambda O_1'O$. 将 $O_1'O$ 和 Λ 一样分块，即记

$$O_1'O=\begin{pmatrix}A_{11}&\cdots&A_{1s}\\\vdots&&\vdots\\A_{s1}&\cdots&A_{ss}\end{pmatrix}=A_0,$$

于是 $(\lambda_j-\lambda_k)A_{jk}=0, 1\leqslant j,k\leqslant s$. 这证明了 $A_{jk}=0, j\neq k$. 所以

$$O_1'O=\mathrm{diag}(A_{11}^{(n_1)},\cdots,A_{ss}^{(n_s)})=A_0,$$

因此 $A_{11},\cdots,A_{ss}$ 都是实对称方阵. 且

$$S_2=O_1\Lambda^{\frac{1}{2}}O_1'=OA_0'\Lambda^{\frac{1}{2}}A_0O'.$$

由 $\Lambda^{\frac{1}{2}}A_0=A_0\Lambda^{\frac{1}{2}}$, 所以有

$$S_2=OA_0'\Lambda^{\frac{1}{2}}A_0O'=OA_0'A_0\Lambda^{\frac{1}{2}}O'=O\Lambda^{\frac{1}{2}}O'=S_1.$$

这证明了开方的唯一性.

显然从证明可知 $S>0$ 当且仅当 $S_1>0$. 最后证明若 $AS=SA$, 则 $AS_1=S_1A$. 事实上, 由 $AS=SA$, 有 $AO\Lambda O'=O\Lambda O'A$, 即 $(O'AO)\Lambda=\Lambda(O'AO)$. 将 $O'AO$ 和 Λ 一样分块, 即记作

$$O'AO=\begin{pmatrix}B_{11} & \cdots & B_{1s}\\ \vdots & & \vdots\\ B_{s1} & \cdots & B_{ss}\end{pmatrix}.$$

于是 $(\lambda_j-\lambda_k)B_{jk}=0$. 因此当 $j\neq k$, 有 $B_{jk}=0$, 所以

$$O'AO=\operatorname{diag}(B_{11}^{(n_1)},\cdots,B_{ss}^{(n_s)}).$$

于是 $O'AO\Lambda^{\frac{1}{2}}=\Lambda^{\frac{1}{2}}O'AO$, 此即 $AO\Lambda^{\frac{1}{2}}O'=O\Lambda^{\frac{1}{2}}O'A$, 因此 $AS_1=S_1A$. 证完.

定理 10.2.5 (极分解式) 设 A 为 n 阶非异实方阵, 则唯一存在 n 阶实正交方阵 O 及实定正对称方阵 S_1 及 S_2, 使得

$$A=S_1O=OS_2$$

这称为 n 阶非异方阵 A 的**极分解**

证 今 A 非异, 所以 $A'A>0$. 因此存在实定正对称方阵 S_2, 使得 $S_2^2=A'A$. 记 $O=AS_2^{-1}$, 所以 $A=OS_2$, 且 $O'O=S_2^{-1}A'AS_2^{-1}=S_2^{-1}(S_2^2)S_2^{-1}=I$, 即 O 为实正交方阵. 再

$$A=OS_2=(OS_2O')O=S_1O,$$

其中 OS_2O' 仍为定正对称方阵. 这证明了极分解的存在性.

下面证唯一性. 今若 $A=OS_2=O_1S_3$, 其中 O,O_1 为实正交方阵, S_2,S_3 为定正对称方阵, 于是有实正交方阵

$$O'O_1=S_2S_3^{-1}.$$

因此$(S_2S_3^{-1})(S_2S_3^{-1})'=S_2S_3^{-2}S_2=I$,此即$S_3^{-2}=S_2^{-2}$，即有$S_2^2=S_3^2$,其中 $S_2>0,S_3>0$,由定理10.2.4的唯一性,所以$S_3=S_2$,因此$O_1=O$. 这证明了极分解 $A=OS_2$ 的唯一性. 考虑 A' 的极分解 $A'=O'S_1$ 之唯一性,便证明了极分解 $A=S_1O$ 的唯一性,证完.

定理 10.2.6 设 A 为 n 阶实非异方阵,则存在 n 阶实正交方阵 O_1,O_2,使得

$$O_1AO_2=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n),$$

其中

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n>0,$$

又 $\lambda_1^2,\lambda_2^2,\cdots,\lambda_n^2$ 为定正对称方阵 AA' 的所有特征根.

证 由定理 10.2.5 的极分解式,有 $A=SO$,其中 S 为 n 阶实定正对称方阵，O 为 n 阶实正交方阵. 由 $S>0$,所以存在 n 阶实正交方阵 O_1,使得

$$S=O_1'\mathrm{diag}(\lambda_1,\cdots,\lambda_n)O_1,$$

其中 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n>0$. 代入,有

$$A=SO=O_1'\mathrm{diag}(\lambda_1,\cdots,\lambda_n)O_1O\ .$$

取 $O_2=O'O_1'$,便证明了 $O_1AO_2=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$.

又 $AA'=SOO'S=S^2=O_1'\mathrm{diag}(\lambda_1^2,\cdots,\lambda_n^2)O_1$, 这证明了 $\lambda_1^2,\lambda_2^2,\cdots,\lambda_n^2$ 为 AA' 的特征根,证完.

定义 两个 $n\times m$ 实矩阵 A 和 B 称为**实正交相抵的**,如果存在 n 阶实正交方阵 O_1 和 m 阶实正交方阵 O_2,使得

$$B=O_1AO_2$$

引理10.2.1 $n\times m$ 实矩阵的实正交相抵为等价关系,

证 显然 $A=I^{(n)}AI^{(m)}$,即 A 和 A 实正交相抵,所以证明了反身性. 又设 $B=O_1AO_2$,则 $A=O_1'BO_2'$, 即若 A 和 B 实正交相抵,则 B 和 A 也实正交相抵,这证明了对称性. 最后,若 $B=O_1AO_2$,

$C=O_3BO_4$，所以

$$C=O_3BO_4=O_3O_1AO_2O_4.$$

这证明了若 A 和 B 实正交相抵，B 和 C 实正交相抵，所以 A 和 C 实正交相抵，即传递性成立. 证完.

定理 10.2.7 设 A 为 $n\times m$ 实矩阵，则存在 n 阶实正交方阵 O_1 和 m 阶实正交方阵 O_2，使得

$$O_1AO_2=\begin{pmatrix}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{pmatrix} & 0 \\ 0 & 0\end{pmatrix},$$

其中 $\lambda_1\geqslant\cdots\geqslant\lambda_r>0$，$r=\mathrm{rank}(A)$. 又 $\lambda_1^2,\cdots,\lambda_r^2$ 为 AA' 的所有非零特征根. 所以 AA' 的非零特征根及 A 的秩为 A 在实正交相抵下的全系不变量.

证 考虑齐次线性方程组 $Ax=0$，它的解空间为 V_m 中的 $m-\mathrm{rank}(A)$ 维子空间，在其中取定标准正交基 $\beta_{r+1},\cdots,\beta_m$，于是在 V_m 中有标准正交基 $\beta_1,\cdots,\beta_r,\beta_{r+1},\cdots\beta_m$，构作 m 阶实正交方阵

$$O_4=(\beta_1\beta_2\cdots\beta_m),$$

则由 $A\beta_{r+1}=0,\cdots,A\beta_m=0$，有

$$AO_4=(A\beta_1\ A\beta_2\cdots A\beta_r\ \ 0\cdots 0)=(A_1^{(n,r)}\ 0^{(n,m-r)}).$$

令 $r=\mathrm{rank}(A)=\mathrm{rank}(AO_4)=\mathrm{rank}(A_1)$. 对 $r\times n$ 实矩阵 A_1' 同上讨论，所以存在 n 阶实正交方阵 O_3，使得

$$A_1'O_3'=(A^{(r)\prime}\ 0^{(r,n-r)}).$$

由 $r=\mathrm{rank}(A_1)=\mathrm{rank}(A_1'O_3')=\mathrm{rank}(A_2)$，所以 A_2 为 r 阶实非异方阵. 由定理 10.2.6，存在 r 阶实正交方阵 O_5,O_6，使得

$$O_5A_2O_6=\mathrm{diag}(\lambda_1,\cdots,\lambda_r),\quad \lambda_1\geqslant\cdots\geqslant\lambda_r>0.$$

于是

$$\begin{pmatrix} O_5 & 0 \\ 0 & I^{(n-r)} \end{pmatrix} O_3 A O_4 \begin{pmatrix} O_6 & 0 \\ 0 & I^{(m-r)} \end{pmatrix} = \begin{pmatrix} \mathrm{diag}(\lambda_1, \cdots, \lambda_r) & 0 \\ 0 & 0 \end{pmatrix}$$

这证明了存在实正交方阵 $O_1^{(n)}$, $O_2^{(m)}$, 使得

$$O_1 A O_2 = \begin{pmatrix} \mathrm{diag}(\lambda_1, \cdots, \lambda_r) & 0 \\ 0 & 0 \end{pmatrix}$$

于是

$$AA' = O_1' \begin{pmatrix} \mathrm{diag}(\lambda_1, \cdots, \lambda_r) & 0 \\ 0 & 0 \end{pmatrix} O_2' O_2 \begin{pmatrix} \mathrm{diag}(\lambda_1, \cdots, \lambda_r) & 0 \\ 0 & 0 \end{pmatrix} O_1$$

$$= O_1' \begin{pmatrix} \mathrm{diag}(\lambda_1^2, \cdots, \lambda_r^2) & 0 \\ 0 & 0 \end{pmatrix} O_1.$$

这证明了 $\lambda_1^2, \cdots, \lambda_r^2$ 为 AA' 的所有非零特征根. 因此 $n \times m$ 实矩阵 A 的秩及 AA' 的所有非零特征根构成 A 在实正交相抵下的全系不变量. 证完.

推论 设 A 为 $n \times m$ 实矩阵, 则

$$\mathrm{rank}(AA') = \mathrm{rank}(A'A) = \mathrm{rank}(A).$$

且 AA' 和 $A'A$ 的非零特征根相同.

证 这是定理 10.2.7 的自然推论.

定理 10.2.8(极分解式) 设 A 为 n 阶实方阵, 则存在 n 阶实正交方阵 O 和 n 阶半定正对称方阵 S_1 和 S_2, 使得

$$A = S_1 O = O S_2.$$

称为 n 阶实方阵 A 的**极分解式**.

证 由定理 10.2.7, 存在 n 阶实正交方阵 O_1 和 O_2, 使得

$$A = O_1 \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} O_2,$$

其中

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0.$$

于是

$$A=(O_1\mathrm{diag}(\lambda_1,\cdots,\lambda_n)O_1')(O_1O_2)$$
$$=(O_1O_2)(O_2'\mathrm{diag}(\lambda_1,\cdots,\lambda_n)O_2).$$

取 $O=O_1O_2, S_1=O_1\mathrm{diag}(\lambda_1,\cdots,\lambda_n)O_1', S_2=O_2'\mathrm{diag}(\lambda_1,\cdots,\lambda_n)O_2$，则 S_1,S_2 半定正，且 O 为实正交方阵．这证明了 A 的极分解式．证完．

习题 10.2

1. 试写出并证明定负及半定负的必要且充分条件．

2. 试求 n 元实函数

$$f(x_1,\cdots,x_n)=\sum_{i,j=1}^{n}a_{ij}x_ix_j+2\sum_{i=1}^{n}b_ix_i+a$$

的极值，其中 n 阶方阵 $S=(a_{ij})$ 为实定正对称方阵．

3. 试证：实对称方阵 $S>0$ 当且仅当 $S^{-1}>0$．

4. 设 S 为实对称方阵．试证：S 的特征根在区间 $[\mu,\nu]$ 中当且仅当 $-\mu I^{(n)}+S\geqslant 0, \nu I^{(n)}-S\geqslant 0$．

5. 设实对称方阵 S_1 和 S_2 的特征根分别落在闭区间 $[a,b]$ 和 $[c,d]$ 中．试证：实对称方阵 S_1+S_2 的特征根落在闭区间 $[a+c,b+d]$ 中．

6. 设 S_1 和 S_2 为 n 阶实半定正对称方阵，且 $S_1S_2=S_2S_1$，试证：S_1S 也是半定正实对称方阵．

7. 设 S 为实定正对称方阵．试证：存在 n 阶实非异上三角方阵 Q，使得 $S=Q'Q$．且若 $S=Q'Q=Q_1'Q_1$，其中 Q_1 也是 n 阶实非异上三角方阵．试证：存在对角方阵 $\mathrm{diag}(\delta_1,\cdots,\delta_n)$，其中 $\delta_j=\pm 1$，$1\leqslant j\leqslant n$，且

$$Q_1=\mathrm{diag}(\delta_1,\cdots,\delta_n)Q.$$

8. 试证：实方阵 A 为非异规范方阵当且仅当 A 有极分解 $A=OS=SO$，其中 S 为实定正对称方阵，O 为实正交方阵．

9. 设 n 阶实对称方阵 S 的前 $n-1$ 个顺序主子式大于零，但是 $\det S=0$．试证：$S\geqslant 0$．

10. 设 $S=(a_{ij})$ 为 n 阶实定正对称方阵．试证：

$$\det S\leqslant a_{11}a_{22}\cdots a_{nn}.$$

且等式成立当且仅当S为对角方阵.

11. (Hadamard 不等式)试证：设$A=(a_{ij})$为n阶实非异方阵，则有

$$|\det A|\leqslant\left(\sum_{j=1}^{n}a_{j1}^2\right)^{\frac{1}{2}}\cdots\left(\sum_{j=1}^{n}a_{jn}^2\right)^{\frac{1}{2}}$$

又等号成立当且仅当A的n个列向量两两正交. 它的几何意义为何?

12. 设S和T为两个n阶实半定正对称方阵，则$\det(S+T)\geqslant\det S$. 当S和T为两个n阶实定正对称方阵，则$\det(S+T)>\det S$

13. 设S为n阶实半定正对称方阵，则有

$$\det S\leqslant S\binom{12\cdots r}{12\cdots r}S\binom{r+1,r+2,\cdots n}{r+1,r+2,\cdots n}.$$

当$S>0$，则等式对$r=1,2,\cdots,n$都成立当且仅当S为对角方阵.

14. 设S_1和S_2都是实对称方阵，其中$S_1\geqslant0$，$\det(S_1+\sqrt{-1}S_2)=0$. 试证：存在$n\times1$非零实矩阵α使得$S_1\alpha=0,S_2\alpha=0$.

15. 设A和B都是n阶实定正对称方阵，P为n阶实方阵. 试证：$A-P'B^{-1}P>0$当且仅当$B-PA^{-1}P'>0$.

16. 设S为n阶实定正对称方阵. 试证：对$n\times1$实矩阵x和y，它有

$$(x'Sy)^2\leqslant(x'Sx)(y'Sy).$$

且等式成立当且仅当x和y线性相关.

17. 试证函数

$$f(x_1,\cdots,x_n)=\sum_{j=1}^{n}x_j^2-\sum_{j=1}^{n-1}x_jx_{j+1}\geqslant0,$$

且求它的极小值，及在限制条件$x_n=1$下的极小值.

18. 设S为n阶实定正对称方阵. 作方阵序列

$$X_0=I,\ X_{k+1}=\frac{1}{2}(X_k+SX_k^{-1}),k=0,1,2,\cdots$$

试证：$\lim\limits_{k\to\infty}X_k$存在，记作$B$. 则$B$为$n$阶实定正对称方阵，且$B^2=S$. 又由$SP=PS$可推出$BP=PB$，其中$P$为$n$阶实方阵. ($\lim\limits_{k\to\infty}X_k=B$定义为：记$X_k=(x_{ij}^{(k)})$，则$\lim\limits_{k\to\infty}x_{ij}^{(k)}=b_{ij}$，显然$B=(b_{ij})$仍为$n$阶方阵).

19. 设A为n阶实定正对称方阵，B为$n\times m$实矩阵. 设$\operatorname{rank}(B)=m$，试求实方阵

$$\begin{pmatrix}A^{(n)} & B^{(n\cdot m)}\\ B' & 0^{(m)}\end{pmatrix}$$

的逆方阵.

20. 试证: $\text{rank}(A'AB)=\text{rank}(AB)$,其中 A 为 $m\times n$ 实矩阵, B 为 $n\times p$ 实矩阵.

21. 设 A 为 n 阶实定正对称方阵,试证:

$$A+A^{-1}-2I^{(n)}\geqslant 0.$$

22. 设 A 为 n 阶实定正对称方阵,试证:

$$(Ax,x)+(A^{-1}x,x)\geqslant 2(x,x).$$

何时等式成立?

23. 设 A 为 n 阶实方阵, $S=\frac{1}{2}(A+A')$. 设 $S>0$. 试证:

$$\det A\geqslant\det S.$$

24. 设 A 为 n 阶半定正对称方阵,且 $I-A\geqslant 0$. 试证: 任取 n 阶实正交方阵 O,则有

$$\det(I-AO)\geqslant\det(I-A).$$

25. 设 $S=(a_{ij})$ 为 n 阶实定正对称方阵,试证:

$$\sum_{i,j\neq k}\left(\det\begin{pmatrix}a_{kk} & a_{ik}\\ a_{kj} & a_{ij}\end{pmatrix}\right)x_ix_j$$

为定正二次型, $k=1,2,\cdots,n$.

26. 设 A 和 B 为 n 阶实对称方阵. 试证: $AB=0$ 当且仅当 $\lambda^n\det(\lambda I^{(n)}-A-B)=\det(\lambda I-A)\det(\lambda I-B)$.

§10.3 实斜对称方阵在实相合下的标准形

定理 10.3.1 设 K 为 n 阶实斜对称方阵,则存在由 K 中元素的有理分式(即两个多元多项式之商)为元素的 n 阶实方阵 P,使得 $\det P=1$,且有

$$P'KP=\text{diag}\left(\begin{pmatrix}0 & a_1\\ -a_1 & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & a_s\\ -a_s & 0\end{pmatrix},0,\cdots,0\right),$$

其中 $a_1,\cdots,a_s$ 为 K 中元素的非零有理分式,且 $|a_1|\geqslant|a_2|\geqslant\cdots\geqslant|a_s|>0$. 因此任一 n 阶实斜对称方阵 K 必实相合于标准形

$$\text{diag}\left(\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},(0,\cdots,0)\right),$$

其中 $2s=\mathrm{rank}(K)$，因此 n 阶实斜对称方阵的秩必为偶数.

证 对阶数 n 作归纳法，今 $K'=-K\neq0$. 用适当的置换方阵作实相合，从而无妨设 $K=(a_{ij})$，$a_{12}\neq0$. 即

$$K=\begin{pmatrix}\begin{pmatrix}0 & a_{12}\\ -a_{12} & 0\end{pmatrix} & A^{(2,\,n-2)}\\ -A' & K_1^{(n-2)}\end{pmatrix},$$

其中 $K_1+K_1'=0$. 作

$$\begin{pmatrix}I^{(2)} & \begin{pmatrix}0 & a_{12}^{-1}\\ -a_{12}^{-1} & 0\end{pmatrix}A\\ 0 & I^{(n-2)}\end{pmatrix}'\begin{pmatrix}\begin{pmatrix}0 & a_{12}\\ -a_{12} & 0\end{pmatrix} & A\\ -A' & K_1\end{pmatrix}$$

$$\cdot\begin{pmatrix}I^{(2)} & \begin{pmatrix}0 & a_{12}^{-1}\\ -a_{12}^{-1} & 0\end{pmatrix}A\\ 0 & I^{(n-2)}\end{pmatrix}$$

$$=\begin{pmatrix}I^{(2)} & 0\\ A'\begin{pmatrix}0 & -a_{12}^{-1}\\ a_{12}^{-1} & 0\end{pmatrix} & I^{(n-2)}\end{pmatrix}$$

$$\cdot\begin{pmatrix}\begin{pmatrix}0 & a_{12}\\ -a_{12} & 0\end{pmatrix} & 0\\ -A' & K_1-A'\begin{pmatrix}0 & a_{12}^{-1}\\ -a_{12}^{-1} & 0\end{pmatrix}A\end{pmatrix}$$

$$=\begin{pmatrix}\begin{pmatrix}0 & a_{12}\\ -a_{12} & 0\end{pmatrix} & 0\\ 0 & K_1-A'\begin{pmatrix}0 & a_{12}^{-1}\\ -a_{12}^{-1} & 0\end{pmatrix}A\end{pmatrix}$$

记

$$K_0=K_1-A'\begin{pmatrix}0 & a_{12}^{-1}\\ -a_{12}^{-1} & 0\end{pmatrix}A$$

则 K_0 为斜对称方阵. 且其中元素由 K 中元素的有理分式构

成.由

$$\begin{pmatrix} I^{(2)} & \begin{pmatrix} 0 & a_{12}^{-1} \\ -a_{12}^{-1} & 0 \end{pmatrix} A \\ 0 & I^{(n-2)} \end{pmatrix}$$

为行列式等于1的方阵，且其中元素为K中元素的有理分式. 又由于有理分式的有理分式仍为有理分式，由归纳法假设，便证明了前一断言，即证明了存在n阶实方阵P_1，其中$\det P_1=1$，且

$$P_1'KP_1=\mathrm{diag}\left(\begin{pmatrix} 0 & a_1 \\ -a_1 & 0 \end{pmatrix},\cdots,\begin{pmatrix} 0 & a_s \\ -a_s & 0 \end{pmatrix},0,\cdots,0\right),$$

其中$|a_1|\geqslant\cdots\geqslant|a_s|>0$. 所以$\mathrm{rank}(K)=\mathrm{rank}(P_1'KP_1)=2s$. 又由

$$\begin{pmatrix} 1 & 0 \\ 0 & a_j^{-1} \end{pmatrix}\begin{pmatrix} 0 & a_j \\ -a_j & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & a_j^{-1} \end{pmatrix}=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

所以完全证明了定理. 证完.

第十一章　酉　空　间

§ 11.1　酉　空　间

在这一章，我们考虑复线性空间 $\mathfrak{L}$. 目的是将第九章和第十章的结果推广到复的情形.

定义　n 维复线性空间 $\mathfrak{L}$ 上定正 Hermite 的 Hermite 双线性函数称为 $\mathfrak{L}$ 上的**内积**.

在 n 维复线性空间 $\mathfrak{L}$ 上取定内积，记作 (α,β). 则有

引理 11.1.1(Cauchy 不等式)　在 n 维复线性空间 $\mathfrak{L}$ 上取定内积 (α,β)，则有

$$|(\alpha,\beta)|^2\leqslant(\alpha,\alpha)(\beta,\beta),\quad \forall\alpha,\beta\in\mathfrak{L}$$

且等式成立当且仅当 α 和 β 线性相关.

证　引进复参数 t，则有 $\alpha+t\beta\in\mathfrak{L}$，于是

$$(\alpha+t\beta,\alpha+t\beta)=(\alpha,\alpha)+\bar{t}(\alpha,\beta)+t\overline{(\alpha,\beta)}+|t|^2(\beta,\beta)\geqslant0,\ \forall t\in C.$$

当 $\beta=0$，不等式 $|(\alpha,\beta)|^2\leqslant(\alpha,\alpha)(\beta,\beta)$ 自然成立. 当 $\beta\neq0$，于是 $(\beta,\beta)>0$. 记 $t=x+\sqrt{-1}y$，$(\alpha,\beta)=u+\sqrt{-1}\,v$，其中 x,y,u,v 实. 于是有

$$(\alpha,\alpha)+2xu+2yv+(x^2+y^2)(\beta,\beta)\geqslant0,$$

此即

$$(\beta,\beta)\left(x+\frac{u}{(\beta,\beta)}\right)^2+(\beta,\beta)\left(y+\frac{v}{(\beta,\beta)}\right)^2+(\alpha,\alpha)-\frac{|(\alpha,\beta)|^2}{(\beta,\beta)}\geqslant0.$$

取 $x=-\frac{u}{(\beta,\beta)}, y=-\frac{v}{(\beta,\beta)}$，便证明了 $|(\alpha,\beta)|^2 \leqslant (\alpha,\alpha)(\beta,\beta)$.

设 $|(\alpha,\beta)|^2=(\alpha,\alpha)(\beta,\beta)$，取 $t=x+\sqrt{-1}y=-\frac{(\alpha,\beta)}{(\beta,\beta)}$，则 $(\alpha+t\beta,\alpha+t\beta)=0$. 由内积定正，所以 $\alpha+t\beta=0$，此即 α 和 β 线性相关. 反之，若 α,β 线性相关，无妨设 $\beta=\lambda_0\alpha$，则 $|(\alpha,\beta)|^2=|\bar{\lambda}_0(\alpha,\alpha)|^2=|\lambda_0|^2(\alpha,\alpha)^2=(\beta,\beta)(\alpha,\alpha)$. 这证明了等式成立. 引理证完.

定义 在 n 维复线性空间 Ω 上取定内积 (α,β)，则数值

$$|\alpha|=\sqrt{(\alpha,\ \alpha)}\geqslant 0$$

称为向量 α 的**长度**. 长度为 1 的向量称为**单位向量**.

显然，长度为零的向量只有零向量.

定义 在 n 维复线性空间 Ω 上取定内积 (α,β). 当 $\alpha=0$ 或 $\beta=0$ 时，我们约定向量 α 和 β 之夹角为直角. 当 $\alpha\neq 0$，$\beta\neq 0$ 时，数值

$$\theta=\cos^{-1}\frac{|(\alpha,\beta)|}{|\alpha||\beta|}$$

称为向量 α 和 β 的**夹角**. 这里约定 $\cos^{-1}$ 取主值，即有

$$0\leqslant\theta\leqslant\pi/2$$

夹角为直角时，即当 $(\alpha,\beta)=0$ 时，我们称向量 α 和 β **互相正交**，记作 $\alpha\perp\beta$.

显然，只有零向量才和任何向量正交.

定义 在 n 维复线性空间 Ω 上取定内积，则 Ω 称为关于内积 (α,β) 的**酉空间**或称为**复 Euclid 空间**.

定义 在 n 维酉空间 Ω 中取定内积 (α,β). Ω 中的基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 称为**标准正交基**，如果

$$(\alpha_i,\alpha_j)=\delta_{ij},\quad 1\leqslant i<j\leqslant n$$

即 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 n 个两两正交的单位向量.

下面和实的 Euclid 空间情形一样，可以给出 Schmidt 正交化，即有

引理 11.1.2 设 $\beta_1, \beta_2, \cdots, \beta_r$ 为 n 维酉空间 Ω 中 r 个两两正交的单位向量，则它们线性无关，所以 $r \leqslant n$. 特别 Ω 中 n 个两两正交的单位向量必为标准正交基.

定理 11.1.1(Schmidt 正交化) 设 Ω 为 n 维酉空间. 在 Ω 中任取一组基 $\{\beta_1, \beta_2, \cdots, \beta_n\}$，则在 Ω 中存在标准正交基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$，使得

$$\begin{aligned}
\alpha_1 &= b_{11}\beta_1, \\
\alpha_2 &= b_{21}\beta_1 + b_{22}\beta_2, \\
&\cdots\cdots\cdots \\
\alpha_n &= b_{n1}\beta_1 + b_{n2}\beta_2 + \cdots + b_{nn}\beta_n,
\end{aligned}$$

且 $b_{11} > 0, \cdots, b_{nn} > 0$.

作为推论，有

定理 11.1.2 在 n 维酉空间 Ω 中任取 r 个两两正交的单位向量 $\alpha_1, \alpha_2, \cdots, \alpha_r$，于是 $r \leqslant n$. 且在 Ω 中存在 $n-r$ 个向量 $\alpha_{r+1}, \cdots, \alpha_n$，使得 $\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n$ 为 Ω 的标准正交基. 特别，任取单位向量 α，则存在一组标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$，使得 α 为指定的第 j 个基向量 α_j.

标准正交基的特点为:

定理 11.1.3 设 Ω 为关于内积 (α, β) 的 n 维酉空间. $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 为 Ω 的标准正交基. 记

$$\alpha = \sum_{j=1}^{n} x_j \alpha_j, \quad \beta = \sum_{k=1}^{n} y_k \alpha_k.$$

则有

$$(\alpha, \beta) = \sum_{j=1}^{n} x_j \bar{y}_j.$$

即在标准正交基下内积的方阵表示为单位方阵.

定义 n 阶复方阵 U 称为**酉方阵**. 如果

$$U\bar{U}'=\bar{U}'U=I^{(n)},$$

即 $\bar{U}'=U^{-1}$, 所有 n 阶酉方阵构成的集合记作 $U(n)$, 称为 **n 阶酉群**.

引理 11.1.3 n 阶酉群有性质

(1) 任取 $U_1,U_2\in U(n)$, 则 $U_1U_2\in U(n)$.

(2) 任取 $U\in U(n)$, 则 $U^{-1}=\bar{U}'\in U(n)$, 又 $\bar{U}U'\in U(n)$.

(3) n 阶单位方阵 $I\in U(n)$.

(4) 任取 $U\in U(n)$, 则 $|\det U|=1$.

定理 11.1.4 设 Ω 为 n 维酉空间, 内积为 (α,β). 任取标准正交基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 再在 Ω 中任取一组基 $\{\beta_1,\beta_2,\cdots,\beta_n\}$, 使得基变换公式为

$$\beta_i=\sum_{j=1}^{n}a_{ji}\alpha_j,\qquad 1\leqslant i\leqslant n.$$

它对应的非异复方阵

$$A=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}.$$

则 $\{\beta_1,\beta_2,\cdots,\beta_n\}$ 为标准正交基当且仅当 A 为酉方阵. 所以在 n 维酉空间 Ω 中取定标准正交基后, 则酉方阵和 Ω 的标准正交基间有一个自然的一一对应关系.

和实正交方阵类似, 酉方阵也有许多有用的性质. 为此, 考虑所有 $n\times 1$ 复矩阵构成的复线性空间 V_n, 在 V_n 中引进内积

$$(x,y)=\left(\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\right)=\sum_{j=1}^{n}x_j\bar{y}_j=x'\bar{y}.$$

显然它是V_n上的定正 Hermite 的 Hermite 双线性函数，且在V_n中有标准正交基

$$e_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\quad e_2=\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix},\quad \cdots,\quad e_n=\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}.$$

设$\mathfrak{L}$为n维酉空间，它有标准正交基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 于是$\mathfrak{L}$中任两向量$\alpha,\beta$，则

$$\alpha=\sum_{j=1}^{n}x_j\alpha_j\longrightarrow\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=x\in V_n,$$

$$\beta=\sum_{j=1}^{n}y_j\alpha_j\longrightarrow\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=y\in V_n.$$

且$\mathfrak{L}$中内积(α,β)有

$$(\alpha,\beta)=\sum_{j=1}^{n}x_j\bar{y}_j=(x,y).$$

自然有对应$\alpha_i\longrightarrow e_i$, $i=1,2,\cdots,n$.

定理 11.1.5 n阶酉方阵有性质

(1) U为n阶酉方阵当且仅当它的n个行向量为V_n中标准正交基；当且仅当它的n个列向量也为V_n中标准正交基.

(2) 任给$n\times m$复矩阵U_1，其中$m\leqslant n$. 设

$$\overline{U}_1'U_1=I^{(m)},$$

则存在$n\times(n-m)$复矩阵U_2，使得n阶复方阵

$$U=(U_1U_2)$$

为 n 阶酉方阵.反之，n 阶酉方阵的前 m 列构成的 $n\times m$ 复矩阵 U_1 适合 $\overline{U}_1'U_1=I^{(m)}$.

特别，任一单位列向量能成为一个 n 阶酉方阵的指定的列向量，所以能成为一个 n 阶酉方阵的第一个列向量.

(3) 任给 $m\times n$ 复矩阵 U_1，其中 $m<n$. 设

$$U_1\overline{U}_1'=I^{(m)}.$$

则存在 $(n-m)\times m$ 复矩阵 U_2，使得 n 阶复方阵

$$U=\begin{pmatrix}U_1\\U_2\end{pmatrix}$$

为 n 阶酉方阵，反之，n 阶酉方阵的前 m 行构成的 $m\times n$ 复矩阵 U_1 适合 $U_1\overline{U}_1'=I^{(m)}$.

特别，任一单位行向量能成为一个 n 阶酉方阵的指定的行向量，所以能成为一个 n 阶酉方阵的第一个行向量.

定理 11.1.6 任给 n 阶复方阵 A，则存在 n 阶酉方阵 U_1，U_2 及 n 阶上三角方阵 T_1，n 阶下三角方阵 T_2，使得 T_1 及 T_2 的对角元素都非负，又有

$$A=U_1T_1=T_2U_2.$$

当 A 为 n 阶非异复方阵时，上述分解唯一.

定义 在 n 维酉空间 $\mathfrak{L}$ 上保持向量长度不变的线性变换称为酉变换.

定理 11.1.7 酉空间 $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 为酉变换当且仅当它保持内积不变.

定理 11.1.8 n 维酉空间 $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 为酉变换当且仅当它在标准正交基下的方阵表示为 n 阶酉方阵，所以在酉空间 $\mathfrak{L}$ 中取定标准正交基后，则 $\mathfrak{L}$ 上酉变换和 n 阶酉方阵间有一个自然的一一对应.

定义 设$\mathfrak{S}$为n维酉空间$\mathfrak{L}$中子集. 记

$$\mathfrak{S}^{\perp}=\{\alpha\in\mathfrak{L}\mid(\alpha,\mathfrak{S})=0\}.$$

则$\mathfrak{S}^{\perp}$称为$\mathfrak{S}$的**正交补**.

定理 11.1.9 n维酉空间$\mathfrak{L}$中子集$\mathfrak{S}$的正交补$\mathfrak{S}^{\perp}$为$\mathfrak{L}$的的子空间,且$\mathfrak{S}\subset(\mathfrak{S}^{\perp})^{\perp}$. 又

$$\mathfrak{L}=\mathfrak{S}^{\perp}\oplus(\mathfrak{S}^{\perp})^{\perp}$$

为空间直接和. 特别,当$\mathfrak{S}$本身为子空间时,则有

$$(\mathfrak{S}^{\perp})^{\perp}=\mathfrak{S}.$$

所以$\mathfrak{L}$有空间直接和分解$\mathfrak{L}=\mathfrak{S}\oplus\mathfrak{S}^{\perp}$.

定义 对n维酉空间$\mathfrak{L}$的子空间$\mathfrak{S}$, 任取$\alpha\in\mathfrak{L}$, 则唯一存在$\beta\in\mathfrak{S},\gamma\in\mathfrak{S}^{\perp}$,使得$\alpha=\beta+\gamma$. 于是$\mathfrak{L}$到$\mathfrak{S}$上的对应$\alpha\to\beta$为$\mathfrak{L}$上线性变换,称为$\mathfrak{L}$到$\mathfrak{S}$上的**正交投影**.

定理 11.1.10 设$\mathscr{A}$为n维酉空间$\mathfrak{L}$上线性变换, 记$\mathscr{A}$的象空间$\mathscr{A}(\mathfrak{L})=\mathfrak{S}$,则$\mathscr{A}$为$\mathfrak{L}$到$\mathfrak{S}$上的正交投影当且仅当

$$(\mathscr{A}(\mathfrak{L}),\ (id-\mathscr{A})(\mathfrak{L}))=0.$$

这时有$\mathscr{A}^2=\mathscr{A}$.

例 1 设A为$2n$阶复方阵. 记

$$J=\begin{pmatrix}0 & I^{(n)}\\ -I^{(n)} & 0\end{pmatrix}.$$

如果A适合条件

$$AJA'=J.$$

试证: $\det A=1$.

证 将A按前n行,前n列分为四块

$$A=\begin{pmatrix}B & C\\ D & E\end{pmatrix}.$$

则条件$AJA'=J$可详细写为

$$BC'=CB',ED'=DE',BE'-CD'=I^{(n)}.$$

下面用摄动法来证明 $\det A=1$. 为此，先证明矩阵方程

$$XC'=CX'$$

必有解 X，其中 X 为 n 阶非异复方阵. 事实上，存在 n 阶非异复方阵 P 及 Q，使得

$$C=P\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q,$$

其中 $r=\operatorname{rank}(C)$. 代入方程，有

$$XQ'\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}P'=P\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}QX'.$$

此即

$$(P^{-1}XQ')\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}(P^{-1}XQ')'.$$

所以 $Y=P^{-1}XQ'$ 为 n 阶非异复方阵，且必须适合

$$Y\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}Y'.$$

显然取 $Y=I$ 即可. 这时，$X=P(Q^{-1})'$，它非异，且有 $XC'=CX'$. 这证明了存在所需要的解 X.

现在引进实参数 $\varepsilon>0$. 作 $2n$ 阶复方阵

$$A_\varepsilon=\begin{pmatrix} B+\varepsilon X & C \\ D & E \end{pmatrix}.$$

则

$$\begin{aligned}\det(B+\varepsilon X)&=\det(BX^{-1}+\varepsilon I)\det X\\&=(\det X)\varepsilon^n+\cdots\end{aligned}$$

为 ε 的 n 次复多项式(这里用到 $\det X\neq 0$)，所以至多 n 个实根. 因此存在无限序列 $\{\varepsilon_j\}$，使得

$$\varepsilon_1>\varepsilon_2>\cdots>\varepsilon_j>\cdots>0,$$

且 $\lim\limits_{j\to\infty}\varepsilon_j=0$, $\det(B+\varepsilon_j X)\neq 0$, $j=1,2,\cdots$.

令

$$\det A_{\varepsilon_j}=\det\begin{pmatrix}B+\varepsilon_j X & C\\ D & E\end{pmatrix}$$

$$=\det\begin{pmatrix}I & 0\\ -D(B+\varepsilon_j X)^{-1} & I\end{pmatrix}\begin{pmatrix}B+\varepsilon_j X & C\\ D & E\end{pmatrix}$$

$$=\det\begin{pmatrix}B+\varepsilon_j X & C\\ 0 & E-D(B+\varepsilon_j X)^{-1}C\end{pmatrix}$$

注意到 $BC'=CB'$, $XC'=CX'$, 所以

$$(B+\varepsilon_j X)C'=C(B+\varepsilon_j X)'.$$

因为 $\det(B+\varepsilon_j X)\neq 0$, 所以有

$$(B+\varepsilon_j X)^{-1}C=C'(B'+\varepsilon_j X')^{-1}.$$

代入，有

$$\det A_{\varepsilon_j}=\det(B'+\varepsilon_j X')\det(E-DC'(B'+\varepsilon_j X')^{-1})$$

$$=\det[E(B'+\varepsilon_j X')-DC']$$

$$=\det(EB'-DC'+\varepsilon_j EX')$$

$$=\det(I+\varepsilon_j EX').$$

双方取 $j\to\infty$, 便有 $\det A=\det I=1$. 证完.

注意，如果限制 A 为 $2n$ 阶实方阵，自然也有 $\det A=1$.

下面在 A 为 $2n$ 阶实方阵的情形，给出比上面更简单的证明. 特点是用复方阵的技巧来处理实方阵.

例 2 设 A 为 $2n$ 阶实方阵，且有 $AJA'=J$, 则 $\det A=1$.

证法一 令

$$A_0=AJ+JA=\begin{pmatrix}B & C\\ D & E\end{pmatrix}\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}+\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}\begin{pmatrix}B & C\\ D & E\end{pmatrix},$$

即

$$A_0=\begin{pmatrix}D-C & B+E\\ -B-E & D-C\end{pmatrix}=\begin{pmatrix}P & Q\\ -Q & P\end{pmatrix}.$$

考虑 $2n$ 阶酉方阵

$$U=\frac{1}{\sqrt{2}}\begin{pmatrix} I^{(n)} & \sqrt{-1}I \\ I^{(n)} & -\sqrt{-1}I \end{pmatrix},$$

则由 P,Q 为 n 阶实方阵，又

$$UA_0\bar{U}'=\frac{1}{2}\begin{pmatrix} P-\sqrt{-1}Q & Q+\sqrt{-1}P \\ P+\sqrt{-1}Q & Q-\sqrt{-1}P \end{pmatrix}\begin{pmatrix} I & I \\ \sqrt{-1}I & \sqrt{-1}I \end{pmatrix}$$

$$=\begin{pmatrix} P-\sqrt{-1}Q & 0 \\ 0 & P+\sqrt{-1}Q \end{pmatrix}$$

所以

$$\det A_0=\det UA_0\bar{U}'=|\det(P+\sqrt{-1}Q)|^2\geqslant 0.$$

而

$$A_0A'=(AJ+JA)A'=AJA'+JAA'=J(I+AA').$$

双方取行列式，有

$$\det A_0\det A=\det(I+AA')\det J=\det(I+AA')>0$$

这是因为 $\det J=1, AA'\geqslant 0$，所以 $I+AA'>0$. 这证明了 $\det A_0>0$，且 $\det A=(\det A_0)^{-1}\det(I+AA')>0$. 但是由 $AJA'=J$，所以 $(\det A)^2=1$. 这证明了 $\det A=1$. 证完.

证法二 令 U 为如上 $2n$ 阶酉方阵，而

$$UA\bar{U}'=\frac{1}{2}\begin{pmatrix} I & \sqrt{-1}I \\ I & -\sqrt{-1}I \end{pmatrix}\begin{pmatrix} B & C \\ D & E \end{pmatrix}\begin{pmatrix} I & I \\ -\sqrt{-1}I & \sqrt{-1}I \end{pmatrix}$$

$$=\frac{1}{2}\begin{pmatrix} B+\sqrt{-1}D & C+\sqrt{-1}E \\ B-\sqrt{-1}D & C-\sqrt{-1}E \end{pmatrix}\begin{pmatrix} I & I \\ -\sqrt{-1}I & \sqrt{-1}I \end{pmatrix}$$

$$=\frac{1}{2}\begin{pmatrix} B+E+\sqrt{-1}(D-C) & B-E+\sqrt{-1}(D+C) \\ B-E-\sqrt{-1}(D+C) & B+E-\sqrt{-1}(D-C) \end{pmatrix}.$$

由于 B,C,D,E 为 n 阶实方阵，所以有

$$T=UA\bar{U}'=\begin{pmatrix}P & Q\\ \bar{Q} & \bar{P}\end{pmatrix},$$

其中

$$2P=B+E+\sqrt{-1}(D-C),\quad 2Q=B-E+\sqrt{-1}(D+C).$$

由 $A=\bar{U}TU$, $AJA'=J$, 所以有 $\bar{U}TUJ\bar{U}'T'U=J$, 即有 $T(UJ\bar{U}')\bar{T}'=(UJ\bar{U}')$. 而

$$UJ\bar{U}'=\begin{pmatrix}-\sqrt{-1}I & 0\\ 0 & \sqrt{-1}I\end{pmatrix}=-\sqrt{-1}\begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix}.$$

所以有

$$\begin{pmatrix}P & Q\\ \bar{Q} & \bar{P}\end{pmatrix}\begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix}\begin{pmatrix}\bar{P}' & Q'\\ \bar{Q}' & P'\end{pmatrix}=\begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix}.$$

此即

$$P\bar{P}'-Q\bar{Q}'=I,\ PQ'=QP'.$$

于是 $P\bar{P}'=I+Q\bar{Q}'>0$, 即有 $\det P\neq 0$, 而

$$\det A=\det T=\det\begin{pmatrix}P & Q\\ \bar{Q} & \bar{P}\end{pmatrix}=\det\begin{pmatrix}I & 0\\ -\bar{Q}P^{-1} & I\end{pmatrix}\begin{pmatrix}P & Q\\ \bar{Q} & \bar{P}\end{pmatrix}$$

$$=\det\begin{pmatrix}P & Q\\ 0 & \bar{P}-\bar{Q}P^{-1}Q\end{pmatrix}=\det P\ \det(\bar{P}-\bar{Q}P^{-1}Q).$$

代入有

$$\det A=\det(\bar{P}P'-\bar{Q}P^{-1}QP')=\det(\bar{P}P'-\bar{Q}Q')=\det I=1.$$

证完.

下面一个例子也是一种标准技巧.

例三 设 A 为 $2n$ 阶酉方阵, 且有 $AJA'=J$, 其中 $J=\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}$, 则 A 称为**辛方阵**. 所有 $2n$ 阶辛方阵构成集合 $Sp(n)$. 试证:

(1) $\forall A,B\in Sp(n)$, 则 $AB, A^{-1}\in Sp(n)$.

(2) 任取 $A\in Sp(n)$, 则存在 $P\in Sp(n)$, 使得

$$\bar{P}'AP=\mathrm{diag}(e^{i\theta_1},\cdots,e^{i\theta_n},e^{-i\theta_1},\cdots,e^{-i\theta_n}).$$

其中 $i=\sqrt{-1}$.特别 $\det A=1$.

证 由定义,(1)显然成立,下面证(2).

设 $A\in Sp(n)$,即有 $A\bar{A}'=I^{(2n)}$, $AJA'=J$. 设 a 为 A 的特征根, α 为 A 的属于特征根 a 的特征向量,且 α 为单位向量,即有

$$A\alpha=a\alpha,\ |\alpha|=1.$$

由 $AJA'=J$,所以 $AJA'\bar{A}=J\bar{A}$,即有 $AJ=J\bar{A}$.记 $\beta=-J\bar{\alpha}$,则有

$$A\beta=-AJ\bar{\alpha}=-J\,\overline{A\alpha}=-\bar{a}J\bar{\alpha}=\bar{a}\beta.$$

又

$$|\beta|=\sqrt{(\beta,\beta)}=\sqrt{\beta'\bar{\beta}}=\sqrt{(J\bar{\alpha})'\,\overline{(J\bar{\alpha})}}=\sqrt{\bar{\alpha}'J'J\alpha}$$
$$=\sqrt{\bar{\alpha}'\alpha}=|\alpha|=1.$$

所以 β 为属于特征根 $\bar{a}$ 的单位特征向量.

为方便起见,仍用 J 及 A 表示 V_{2n} 上线性变换,使得它们在 V_{2n} 中标准基下对应之方阵表示分别为 J 及 A、今

$$(\alpha,\beta)=\alpha,\bar{\beta}=\alpha'(\overline{-J\bar{\alpha}})=-\alpha'J\alpha=0,$$

所以 $\alpha\perp\beta$. 于是 α, β 为 V_{2n} 中 2 维子空间 $\mathfrak{R}_1$ 的标准正交基. 由于

$$V_{2n}=\mathfrak{R}_1\oplus\mathfrak{R}_1^{\perp},$$

下面证明 $\mathfrak{R}_1$ 及 $\mathfrak{R}_1^{\perp}$ 都在 V_{2n} 上线性变换 A 作用下不变. 即有 $A\mathfrak{R}_1\subset\mathfrak{R}_1$, $A\mathfrak{R}_1^{\perp}\subset\mathfrak{R}_1^{\perp}$. 再证明 $J\overline{\mathfrak{R}_1}\subset\mathfrak{R}_1$, $J\overline{\mathfrak{R}_1^{\perp}}\subset\mathfrak{R}_1^{\perp}$. 事实上,任取 $\mathfrak{R}_1$ 中向量 $\xi=c_1\alpha+c_2\beta=c_1\alpha-c_2J\bar{\alpha}$,其中 c_1, c_2 为复数,于是由 $J^2=-I^{(2n)}$, $JJ'=J'J=I^{(2n)}$,有

$$A\xi=c_1A\alpha+c_2A\beta=ac_1\alpha+\bar{a}c_2\beta\in\mathfrak{R}_1$$
$$J\bar{\xi}=J(\bar{c}_1\bar{\alpha}-\bar{c}_2J\alpha)=\bar{c}_1J\bar{\alpha}+\bar{c}_2\alpha=\bar{c}_2\alpha-\bar{c}_1\beta\in\mathfrak{R}_1$$

这证明了 $A\mathfrak{R}_1\subset\mathfrak{R}_1$, $J\overline{\mathfrak{R}_1}\subset\mathfrak{R}_1$.

再任取 $\eta\in\mathfrak{R}_1^{\perp}$, 于是 $(\alpha,\ \eta)=\eta'\bar{\alpha}=0$, $(\beta,\ \eta)=\eta'\bar{\beta}=$

$-\eta'\overline{(J\bar{\alpha})}=-\eta'J\alpha=0$,即有

$$\bar{\eta}'\alpha=0,\ \eta'J\alpha=0.$$

为了证 $A(\mathfrak{N}_1^\perp)\subset\mathfrak{N}_1^\perp$,只要证 $(\overline{A\eta})'\alpha=0$, $(A\eta)'J\alpha=0$. 事实上,由 $\bar{A}'=A^{-1}$ 及 $A^{-1}J=JA'$,所以有

$$(\overline{A\eta})'\alpha=\bar{\eta}'\bar{A}'\alpha=\bar{\eta}'A^{-1}\alpha=\bar{\eta}'\left(\frac{1}{a}\alpha\right)=\frac{1}{a}\bar{\eta}'\alpha=0,$$

$$(A\eta)'J\alpha=\eta'A'J\alpha=\eta'(\overline{A'J})\alpha=\eta'(\overline{JA'})\alpha$$

$$=\eta'JA^{-1}\alpha=\frac{1}{a}\eta'J\alpha=0$$

为了证 $J(\overline{\mathfrak{N}_1^\perp})\subset\mathfrak{N}_1^\perp$,只要证$(\overline{J\bar{\eta}})'\alpha=0$, $(J\bar{\eta})'J\alpha=0$.事实上

$$(\overline{J\bar{\eta}})'\alpha=\eta'J'\alpha=-\eta'J\alpha=0,$$

$$(J\bar{\eta})'J\alpha=\bar{\eta}'J'J\alpha=\bar{\eta}'\alpha=0.$$

至此证明了断言.

现在对维数作归纳法. 因此证明了在 V_{2n} 中存在标准正交基 $\alpha_1,\cdots,\alpha_n,\beta_1=-J\bar{\alpha}_1,\cdots,\beta_n=-J\bar{\alpha}_n$,使得

$$A\alpha_j=a_j\alpha_j,\ A\beta_j=\bar{a}_j\beta_j,\ 1\leqslant j\leqslant n.$$

由于 A 为酉方阵,特征根的模为 1,所以

$$a_j=e^{i\theta_j},\ 1\leqslant j\leqslant n.$$

作 $2n$ 阶酉方阵

$$P=(\alpha_1\cdots\alpha_n\ \ \beta_1\cdots\beta_n)$$

我们来证 $P\in Sp(n)$. 事实上,由 $J^2=-I^{(2n)}$,有

$$PJP'=(\alpha_1\cdots\alpha_n\ \ \beta_0\cdots\beta_n)\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}P'$$

$$=(-\beta_1\cdots-\beta_n\ \ \alpha_1\cdots\alpha_n)P'$$

$$=J(J\beta_1\cdots J\beta_n\ \ -J\alpha_1\ \cdots\ -J\alpha_n)P'$$

$$=J(\bar{\alpha}_1\cdots\bar{\alpha}_n\ \ \bar{\beta}_1\cdots\bar{\beta}_n)P'$$

$$=J\bar{P}P'=J.$$

此即 $P\in Sp(n)$. 再

$$
\begin{aligned}
\bar{P}'AP &= \bar{P}'A(\alpha_1\cdots\alpha_n \;\; \beta_1\cdots\beta_n) \\
&= \bar{P}'(A\alpha_1\cdots A\alpha_n \;\; A\beta_1\cdots A\beta_n) \\
&= \bar{P}'(a_1\alpha_1\cdots a_n\alpha_n \;\; \bar{a}_1\beta_1\cdots\bar{a}_n\beta_n) \\
&= \bar{P}'(\alpha_1\cdots\alpha_n \;\; \beta_1\cdots\beta_n)\operatorname{diag}(a_1,\cdots,a_n,\bar{a}_1,\cdots,\bar{a}_n) \\
&= \operatorname{diag}(a_1,\cdots,a_n,\bar{a}_1,\cdots,\bar{a}_n).
\end{aligned}
$$

这证明了(2)．这时显然有 $\det A=1$. 证完.

习题 11.1

1．习题 9.1 中哪些可以推广到复的情形

2．记 $\omega=\exp\dfrac{2\pi\sqrt{-1}}{n}$，试证：$n$ 阶复方阵

$$
\frac{1}{\sqrt{n}}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \cdots & \omega^{(n-1)^2} \end{pmatrix}
$$

为酉方阵.

3．试证：任一二阶酉方阵 U 必可分解为

$$
U=\begin{pmatrix} e^{i\theta_1} & 0 \\ 0 & e^{i\theta_2} \end{pmatrix}\begin{pmatrix} \cos\tau & \sin\tau \\ -\sin\tau & \cos\tau \end{pmatrix}\begin{pmatrix} e^{i\theta_3} & 0 \\ 0 & e^{i\theta_4} \end{pmatrix},
$$

其中 $\theta_1,\theta_2,\theta_3,\theta_4,\tau$ 都是实数

4．记 $\mathfrak{L}$ 为所有 $n\times m$ 复矩阵构成的 nm 维复线性空间．任取 $n\times m$ 复矩阵 A,B．试证：

$$
(A,B)=\operatorname{tr}A\bar{B}'
$$

为 $\mathfrak{L}$ 上内积．并求出一组标准正交基.

§ 11.2　在酉相似下复方阵的标准形

定义　n 阶复方阵 A 和 B 称为酉相似的，如果存在 n 阶酉方阵 U，使得

$$
B=UAU^{-1}=UA\bar{U}'.
$$

显然酉相似为等价关系.

定理 11.2.1 n阶复方阵A酉相似于下(上)三角方阵

$$B=\begin{pmatrix} a_{11} & & \\ \vdots & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

其中对角元素为复方阵A的所有特征根.

注 由于复方阵的特征根都是复数. 所以这个定理的证明比定理 9.2.1 的证明要简单. 这也可以看出,处理复的情形,往往比处理实的情形简单.

下面进一步考虑一些特殊的方阵在酉相似下的标准形.

定理 11.2.2(Schur 不等式) 记$\lambda_1,\cdots,\lambda_n$为$n$阶复方阵$A$的$n$个特征根,则有不等式

$$\sum_{j=1}^{n}|\lambda_j|^2\leqslant \operatorname{tr}A\overline{A}'.$$

且等号成立当且仅当复方阵A适合条件

$$A\overline{A}'=\overline{A}'A$$

这类方阵称为**复规范方阵**,或简称为**规范方阵**.

注 同样,这个定理的证明比定理 9.2.2 的证明简单.

定理 11.2.3 n阶复规范方阵A酉相似于对角形

$$B=\operatorname{diag}(\lambda_1,\cdots,\lambda_n),$$

其中$\lambda_1,\cdots,\lambda_n$为$A$的所有特征根. 所以规范方阵$A$在酉相似下的全系不变量为$A$的所有特征根.

定理 11.2.4 n阶酉方阵U酉相似于对角形

$$B=\operatorname{diag}(e^{\sqrt{-1}\theta_1},\cdots,e^{\sqrt{-1}\theta_n}),$$

其中$0\leqslant\theta_1\leqslant\cdots\leqslant\theta_n<2\pi$. 所以酉方阵的特征根的模为1,且酉方阵$U$在酉相似下的全系不变量为$U$的所有特征根.

定理 11.2.5 n阶 Hermite 方阵H酉相似于对角形

$$\operatorname{diag}(\lambda_1,\cdots,\lambda_n),$$

其中 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ 都是实数，且为H的所有特征根. 所以Hermite 方阵的特征根都是实数，且 Hermite 方阵在酉相似下的全系不变量为它的所有特征根.

定理 11.2.6 n阶斜 Hermite 方阵K酉相似于对角形

$$\operatorname{diag}(\sqrt{-1}\lambda_1, \cdots, \sqrt{-1}\lambda_n)$$

其中 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ 都是实数，且$\sqrt{-1}, \cdots, \sqrt{-1}\lambda_n$ 为K的所有特征根、所以斜 Hermite 方阵的非零特征根都是纯虚数，且斜 Hermite方阵在酉相似下的全系不变量为它的所有特征根.

引理 11.2.1 设Ω为n维酉空间，$\mathscr{A}$为Ω上线性变换，则它在Ω的不同标准正交基下对应的方阵表示互相酉相似.

定理 11.2.7 设(α, β)为n维酉空间Ω上内积. 任取线性函数f，则在Ω中唯一存在向量α_f，使得$f(\beta) = (\alpha_f, \beta)$，$\forall \beta \in \Omega$. 且对应$f \to \alpha_f$给出了对偶空间$\Omega^*$到$\Omega$上的线性同构.

定理 11.2.8 记(α, β)为n维酉空间Ω上的内积. 设$\mathscr{A}$为Ω上线性变换. 任给$\beta \in \Omega$，则线性函数$f(\alpha) = (\mathscr{A}(\alpha, \beta)$，$\forall \alpha \in \Omega$唯一决定$\Omega$中向量$\beta_f$，它只与$\mathscr{A}$及$\beta$有关，改记为$\mathscr{A}^*(\beta)$，即有

$$(\mathscr{A}(\alpha), \beta) = (\alpha, \mathscr{A}^*(\beta)), \forall \alpha, \beta \in \Omega.$$

则$\beta \to \mathscr{A}^*(\beta)$为$\Omega$上线性变换，记作$\mathscr{A}^*$，称为$\mathscr{A}$的**共轭变换**. 在$n$维酉空间$\Omega$中取定标准正交基$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$，记线性变换$\mathscr{A}$对应的方阵表示为$A$，则共轭变换$\mathscr{A}^*$对应的方阵表示为$\overline{A}'$.

定义 记(α, β)为n维酉空间Ω的内积. Ω上线性变换称为**规范的**(**酉的**、**Hermite 的**、**斜 Hermite 的**)，如果有$\mathscr{A}\mathscr{A}^* = \mathscr{A}^*\mathscr{A}$ ($\mathscr{A}\mathscr{A}^* = \mathscr{A}^*\mathscr{A} = id$、$\mathscr{A}^* = \mathscr{A}$、$\mathscr{A}^* = -\mathscr{A}$)，即有$(\mathscr{A}(\alpha), \mathscr{A}(\beta)) = (\mathscr{A}^*(\alpha), \mathscr{A}^*(\beta))$. ($(\mathscr{A}(\alpha), \mathscr{A}(\beta)) = (\alpha, \beta)$、$(\mathscr{A}(\alpha), \beta) = (\alpha, \mathscr{A}(\beta))$、$(\mathscr{A}(\alpha), \beta) + (\alpha, \mathscr{A}(\beta)) = 0$, $\forall \alpha, \beta \in \Omega$.

所以定理 11.2.3 到定理 11.2.6 可改写为

定理 11.2.9 记(α, β)为n维酉空间Ω的内积，$\mathscr{A}$为Ω上

线性变换．则有$\mathscr{A}$为规范变换当且仅当在$\mathfrak{V}$中存在标准正交基$\{\alpha_1,\cdots,\alpha_n\}$，使得

$$\mathscr{A}(\alpha_j)=\lambda_j\alpha_j,\quad j=1,2,\cdots,n.$$

其中$\lambda_1,\cdots,\lambda_n$为复数．于是$\mathscr{A}$为酉变换当且仅当$\lambda_j=e^{\sqrt{-1}\theta_j},1\leqslant j\leqslant n$；$\mathscr{A}$为Hermite变换当且仅当$\lambda_1,\cdots,\lambda_n$都为实数；$\mathscr{A}$为斜Hermite变换当且仅当$\lambda_1,\cdots\lambda_n$都为纯虚数或零．

习题 11.2

1. 习题9.2中哪些可以推广到复的情形．

2. 设A为n阶规范方阵，它的n个特征根

$$\lambda_j=a_j+\sqrt{-1}b_j,\quad 1\leqslant j\leqslant n,$$

其中$a_1,\cdots,a_n$，b_1，$\cdots$，b_n为$2n$个不同实数．试证：若α为方阵A，$B=\frac{1}{2}(A+\bar{A}')$，$C=\frac{1}{2\sqrt{-1}}(A-\bar{A}')$中某一个方阵的特征向量，则必存在复数$\lambda=\mu+\sqrt{-1}\nu$，其中$\mu,\nu$为实数，使得

$$A\alpha=\lambda\alpha,\ B\alpha=\mu\alpha,\ C\alpha=\nu\alpha.$$

3. 设A为斜Hermite方阵，试证：$U=(I+A)(I-A)^{-1}$为酉方阵，且特征根不等于-1．反之，设n阶酉方阵U的特征根不等于-1，则

$$K=(I-U)(I+U)^{-1}$$

为斜Hermite方阵．这种关系称为Cayley变换．

4. 任取$\alpha\in C^n$．试证：存在n阶实正交方阵O，使得

$$(O\alpha)'=(\lambda_1e^{\sqrt{-1}\theta},-\sqrt{-1}\lambda_2e^{-\sqrt{-1}\theta},0,\cdots,0),$$

$$(O\alpha)'=e^{\sqrt{-1}\theta}(\lambda_1,\sqrt{-1}\lambda_2,0,\cdots,0),$$

其中$\lambda_1\geqslant\lambda_2\geqslant0$．

5. 设A为n阶规范方阵，A的模最大的特征根为λ_1，模最小的特征根为λ_n．设U为n阶酉方阵，且$\det(\lambda_0A-U)=0$．试证：

$$|\lambda_n|\leqslant|\lambda_0|^{-1}\leqslant|\lambda_1|.$$

6. 记A为n阶复方阵，$H=\frac{1}{2}(A+\bar{A}')$，$K=\frac{1}{2}(A=\bar{A}')$．记$a,h,k$分别

为 A,H,K 中元素的最大模. 记 $\lambda=b+\sqrt{-1}c$ 为 A 的特征根，其中 b,c 为实数. 试证：

$$|\lambda|\leqslant na,\quad |b|\leqslant nh,\quad |c|\leqslant nk.$$

以此来证明 Hermite 方阵及斜 Hermite 方阵、酉方阵的特征根分别为实数、纯虚数及模为 1 的复数.

7. 记 A 为 n 阶复方阵, $H=\frac{1}{2}(A+\bar{A}')$, $K=\frac{1}{2}(A-\bar{A}')$. 记 λ_1, λ_2, $\cdots$, λ_n 为 A 的 n 个特征根. 试证：

$$\sum_{j=1}^{n}|\lambda_j|^2\leqslant \mathrm{tr}A\bar{A}',\quad \sum_{j=1}^{n}(\mathrm{Re}(\lambda_j))^2\leqslant \mathrm{tr}H\bar{H}',$$

$$\sum_{j=1}^{n}(\mathrm{Im}(\lambda_j))^2\leqslant \mathrm{tr}K\bar{K}'.$$

且其中有一个等式成立当且仅当 A 为规范方阵，当且仅当三个等式同时成立.

8. 设 n 阶酉方阵 S 为复对称方阵 $S'=S$，则 S 称为**酉对称方阵**. 试证：

(i) S 实正交相似于对角方阵 $\mathrm{diag}(e^{\sqrt{-1}\theta_1},\cdots,e^{\sqrt{-1}\theta_n})$.

(ii) 存在酉方阵 U 使得 $S=UU'$.

(iii) 存在实对称方阵 S_1，使得

$$S=\lim_{k\to\infty}\sum_{l=0}^{k}\frac{(-1)^{l/2}}{l!}S_1^l=\sum_{l=0}^{\infty}\frac{(-1)^{l/2}}{l!}S_1^l.$$

(iv) 任一酉方阵必可分解为一个实正交方阵和一个酉对称方阵的乘积. 试讨论它的唯一性.

(v) 任一酉方阵实正交相抵于对角形.

9. n 阶 Hermite 方阵 O 称为正交 Hermite 的，如果它还是复正交方阵，即有 $OO'=O'O=I$. 试证：

(i) 正交 Hermite 方阵 O 实正交相似于标准形

$$\mathrm{diag}\left(\begin{pmatrix} a_1 & \sqrt{-1}b_1 \\ -\sqrt{-1}b_1 & a_1\end{pmatrix},\cdots,\begin{pmatrix} a_s & \sqrt{-1}b_s \\ -\sqrt{-1}b_s & a_s\end{pmatrix},\right.$$
$$\left.1,\cdots,1,-1,\cdots,-1\right),$$

其中 $a_1,\cdots,a_s,b_1,\cdots,b_s$ 都是实数，且 $b_1\geqslant\cdots\geqslant b_s>0$，又 $a_j=\pm\sqrt{1+b_j^2}$, $1\leqslant$

$j \leqslant s$.

(ii) 正交 Hermite 方阵 O 必可分解为一个实正交对称方阵 A 和一个实斜对称方阵 K 定义的 n 阶方阵

$$e^{\sqrt{-1}K} = \lim_{k \to \infty} \sum_{l=0}^{k} \frac{(-1)^{l/2}}{l!} K^l = \sum_{l=0}^{\infty} \frac{(-1)^{l/2}}{l!} K^l$$

之乘积

$$O = Ae^{\sqrt{-1}K},$$

其中 $AK = KA, A^2 = I^{(n)}$. 试讨论分解的唯一性.

(iii) 任一复正交方阵 O 必可表为

$$O = Ae^{\sqrt{-1}k},$$

其中 A 实正交, K 实斜对称. 试讨论分解的唯一性.

10. 记 $2n$ 阶方阵

$$J = \text{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right).$$

设 $2n$ 阶 Hermite 方阵 H 适合

$$HJ = J\bar{H}.$$

则有

(i) 若 α 为 H 的属于特征根 λ_0 的特征向量, 则 $J\alpha$ 为 Hermite 方阵 H' 的属于特征根 λ_0 的特征向量.

(ii) H 的特征根必为重根, 记作 $\lambda_1, \lambda_1, \cdots, \lambda_n, \lambda_n$, 其中 $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$. 且 H 酉相似于对角形

$$\Lambda = \text{diag}(\lambda_1 I^{(2)}, \cdots, \lambda_n I^{(2)}).$$

又存在酉方阵 U 使得 $H = U\Lambda\bar{U}', UJ = J\bar{U}$.

(iii) $2n$ 阶复方阵 A 若适合 $AJ = J\bar{A}$, 则存在酉方阵 U, 适合 $UJ = J\bar{U}$, 且 $\bar{U}'AU$ 为上三角方阵.

§11.3 Hermite 型的分类

定义 n 阶复方阵 A 和 B 称为**复相合的**, 如果存在 n 阶可逆复方阵 P, 使得

$$B = \bar{P}'AP.$$

引理 11.3.1 复相合为等价关系.

显然 Hermite 方阵（斜 Hermite 方阵）在复相合下仍变为 Hermite 方阵(斜 Hermite 方阵).

定理 11.3.1 任一 Hermite 方阵H在复相合下的标准形为

$$\operatorname{diag}(I^{(p)}, -I^{(r-p)}, 0),$$

其中 $r=\operatorname{rank}(H)$. 数值

$$\delta(H)=p-(r-p)=2p-r$$

称为H的**符号差**,它是标准形的迹,也是H的正特征根个数减负特征根的个数.

引理 11.3.2 (witt 定理) 给定n阶 Hermite 方阵 H 和m阶 Hermite 方阵 H_1 和 H_2. 设 H, H_1, H_2 都非异, 且 $n+m$ 阶 Hermite 方阵

$$\begin{pmatrix} H & 0 \\ 0 & H_1 \end{pmatrix}, \quad \begin{pmatrix} H & 0 \\ 0 & H_2 \end{pmatrix}$$

互相复相合,则 H_1 和 H_2 互相复相合.

定理 11.3.2(惯性定理) 两个 n 阶方阵

$$A=\begin{pmatrix} I^{(p)} & 0 & 0 \\ 0 & -I^{(r-p)} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad B=\begin{pmatrix} I^{(q)} & 0 & 0 \\ 0 & -I^{(s-q)} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

复相合当且仅当 $r=s$, $p=q$. 所以 n 阶 Hermite 方阵在复合下的全系不变量为它的秩和符号差.

运用上面理论,可给出 Hermite 型的标准形. 已知 Hermite 型可表为

$$\varphi(x)=\sum_{i,j=1}^{n} a_{ij}x_i\bar{x}_j=x'H\bar{x},$$

其中

$$H=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

为 Hermite 方阵. 作坐标变换

$$x_i=\sum_{j=1}^{n} p_{ij}y_j, \quad i=1,2,\cdots,n,$$

其中

$$P=(p_{ij})$$

为 n 阶非异复方阵. 于是可写成 $x=Py$. 因此

$$\varphi(x)=x'H\bar{x}=y'(P'H\bar{P})\bar{y}.$$

所以同一 Hermite 型关于坐标 x, y 分别对应 Hermite 方阵为 H 和 $P'H\bar{P}$, 即它们互相复合. 所以定理 11.3.1 给出了

定理 11.3.3 任给 Hermite 型

$$\varphi(x)=\sum_{i,j=1}^{n} a_{ij}x_i\bar{x}_j, \quad a_{ij}=\bar{a}_{ji},$$

则存在 n 个独立未知数 $x_1,\cdots,x_n$ 的非异线性变换

$$x_i=\sum_{j=1}^{n} p_{ij}y_j, \quad i=1,2,\cdots,n$$

使得 $\varphi(x)$ 变为 n 个独立未知数 $y_1,\cdots,y_n$ 的 Hermite 型

$$\psi(y)=\sum_{j=1}^{p}|y_j|^2-\sum_{j=p+1}^{r}|y_j|^2$$

上面定理的证明, 和 §10.1 完全类似. 所以我们略去证明. 例如, 关于配方法, 即矩阵打洞技巧, 可以写出如下: 设 H 为 n 阶 Hermite 方阵. 将它按前 s 行、列分块为

$$H=\begin{pmatrix} H_{11}^{(s)} & H_{12} \\ \bar{H}'_{12} & H_{22} \end{pmatrix}.$$

假设 $\det H_{11}\neq 0$,即 H_{11} 为 s 阶非异 Hermite 方阵,则有

$$\overline{\begin{pmatrix} I & -H_{11}^{-1}H_{12} \\ 0 & I \end{pmatrix}'}\begin{pmatrix} H_{11} & H_{12} \\ \overline{H}'_{12} & H_{22} \end{pmatrix}\begin{pmatrix} I & -H_{11}^{-1}H_{12} \\ 0 & I \end{pmatrix}$$

$$=\begin{pmatrix} I & 0 \\ -\overline{H}'_{12}H_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} H_{11} & 0 \\ \overline{H}'_{12} & H_{22}-\overline{H}'_{12}\overline{H}'^{-1}_{11}H_{12} \end{pmatrix}$$

$$=\begin{pmatrix} H_{11} & 0 \\ 0 & H_{22}-\overline{H}'_{12}H_{11}^{-1}H_{12} \end{pmatrix}.$$

定义 n 维复线性空间 Ω 上 Hermite 双线性函数 h 称为非退化的,如果由 $h(\alpha,\alpha)=0$,可推出 $\alpha=0$.

下面给出非退化 Hermite 双线性函数的标准形. 由 §8.1 可知:在 Ω 中取定基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$, 则 Hermite 双线性函数有坐标表达式

$$h(\alpha,\beta)=x'A\overline{y},$$

其中

$$A=\begin{pmatrix} h(\alpha_1,\alpha_1) & \cdots & h(\alpha_1,\alpha_n) \\ \vdots & & \vdots \\ h(\alpha_n,\alpha_1) & \cdots & h(\alpha_n,\alpha_n) \end{pmatrix},\quad x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\quad y=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

又

$$\alpha=\sum_{j=1}^{n}x_j\alpha_j,\qquad \beta=\sum_{j=1}^{n}y_j\alpha_j,$$

而 A 为 n 阶复方阵. 于是记

$$H_1=\frac{1}{2}(A+\overline{A}'),\quad H_2=\frac{1}{2\sqrt{-1}}(A-\overline{A}')$$

则 H_1,H_2 为 n 阶 Hermite 方阵,且

$$A=H_1+\sqrt{-1}H_2.$$

又在 Ω 中另取一组基 $\{\beta,\cdots,\beta_n\}$. 记基变换公式为

$$\beta_j=\sum_{k=1}^{n}p_{kj}\alpha_k,\quad j=1,2,\cdots,n,$$

则

$$P=\begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

为 n 阶非异复方阵. 记

$$B=\begin{pmatrix} h(\beta_1,\beta_1) & \cdots & h(\beta_1,\beta_n) \\ \vdots & & \vdots \\ h(\beta_n,\beta_1) & \cdots & h(\beta_n,\beta_n) \end{pmatrix},$$

$$u=\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, \qquad v=\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

又

$$\alpha=\sum_{j=1}^{n} u_j\beta_j, \qquad \beta=\sum_{j=1}^{n} v_j\beta_j,$$

则有

$$h(\alpha,\beta)=u'B\bar{v}.$$

于是

$$\frac{1}{2}(B+\bar{B}')=P'\left(\frac{1}{2}(A+\bar{A}')\right)\bar{P},$$

$$\frac{1}{2\sqrt{-1}}(B-\bar{B}')=P'\left(\frac{1}{2\sqrt{-1}}(A-\bar{A}')\right)\bar{P}.$$

因此记 n 阶 Hermite 方阵

$$\widetilde{H}_1=\frac{1}{2}(B+\bar{B}'), \qquad \widetilde{H}_2=\frac{1}{2\sqrt{-1}}(B-\bar{B}').$$

则有

$$\widetilde{H}_1=P'H_1\bar{P}, \quad \widetilde{H}_2=P'H_2\bar{P}.$$

而

$$B=\widetilde{H}_1+\sqrt{-1}\widetilde{H}_2=P'H_1\bar{P}+\sqrt{-1}P'H_2\bar{P}=P'A\bar{P}.$$

由此可见，所谓 Hermite 双线性函数化标准形问题，即为：给定 n 阶 Hermite 方阵偶 (H_1, H_2)，寻找 n 阶非异复方阵 P，使得 $(P'H_1\bar{P}, P'H_2\bar{P})$ 为标准形.

引理 11.3.3 在 n 维复线性空间中取定基，则 Hermite 双线性函数 $h(\alpha,\beta)$ 非退化，当且仅当它决定的 n 阶 Hermite 方阵偶 (H_1, H_2) 有：由

$$x'H_1\bar{x}=0, \quad x'H_2\bar{x}=0$$

可推出 $x=0$.

证 今 $h(\alpha,\beta)=x'H_1\bar{y}+\sqrt{-1}x'H_2\bar{y}$. 由定义，$h(\alpha,\beta)$ 非退化当且仅当由 $h(\alpha,\alpha)=0$ 可推出 $\alpha=0$，即由

$$x'H_1\bar{x}+\sqrt{-1}x'H_2\bar{x}=0$$

可推出 $x=0$. 由于 $x'H_1\bar{x}$, $x'H_2\bar{x}$ 都是实数. 此即由 $x'H_1\bar{x}=0$, $x'H_2\bar{x}=0$ 可推出 $x=0$. 证完.

定理 11.3.4 设 H_1 和 H_2 为 n 阶Hermite方阵. 适合条件：设 $x_0\in C^n$，且由 $x_0'H_1\bar{x}_0=0$, $x_0'H_2\bar{x}_0=0$，则可推出 $x_0=0$. 于是存在 n 阶非异复方阵 P，使得

$$\bar{P}'H_1P=\mathrm{diag}(I^{(p)}, -I^{(r-p)}, 0)$$

$$\bar{P}'H_2P=\mathrm{diag}(\lambda_1, \cdots, \lambda_r, \pm I^{(n-r)})$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_r$ 为实数.

证 由定理 11.3.1，存在 n 阶非异复方阵 P_1，使得

$$\bar{P}_1'H_1P_1=\mathrm{diag}(I^{(p)}, -I^{(r-p)}, 0).$$

记

$$\bar{P}_1'H_2P_1=\begin{pmatrix} A^{(r)} & B \\ \bar{B}' & D \end{pmatrix}, \quad A=\bar{A}', \ D=\bar{D}'.$$

于是任取 n 阶非异复方阵

$$P_2=\mathrm{diag}(I^{(r)}, P_3)$$

则有

$$\bar{P}_2'\bar{P}_1'H_1P_1P_2=\operatorname{diag}(I^{(p)},-I^{(r-p)},0),$$

$$\bar{P}_2'\bar{P}_1'H_2P_1P_2=\begin{pmatrix} A & BP_3 \\ \bar{P}_3'\bar{B}' & \bar{P}_3'DP_3 \end{pmatrix}.$$

由定理 11 3.1，存在 $n-r$ 阶非异复方阵 P_3，使得

$$\bar{P}_3'DP_3=\operatorname{diag}(I^{(q)},-I^{(s-q)},0).$$

于是在复相合意义下．无妨设

$$H_1=\operatorname{diag}(I^{(p)},-I^{(r-p)},0^{(n-r)}),$$

$$H_2=\begin{pmatrix} A^{(r)} & B_1 & B_2 & B_3 \\ \bar{B}_1' & I^{(q)} & 0 & 0 \\ \bar{B}_2' & 0 & -I^{(s-q)} & 0 \\ \bar{B}_3' & 0 & 0 & 0 \end{pmatrix}.$$

设 $s<n-r$，取 $x_0=e_n=(0,\cdots,0,1)'$，则有

$$x_0'H_1\bar{x}_0=0,\qquad x_0'H_2\bar{x}_0=0.$$

这和条件 $x_0=0$ 矛盾．所以证明了 $s=n-r$．设 $1\leqslant q<n-r$，于是 $0<s-q<n-r$．这时任取

$$x_0=\begin{pmatrix} 0^{(r,1)} \\ \alpha^{(q,1)} \\ \beta^{(s-q,1)} \end{pmatrix}\in C^n,$$

则有 $$x_0'H_1\bar{x}_0=0,\quad x_0'H_2\bar{x}_0=\alpha'\bar{\alpha}-\beta'\bar{\beta}$$

取 $\alpha=(1,0,\cdots,0)'\in C^q$，$\beta=(1,0,\cdots,0)'\in C^{s-q}$，则有

$$x_0'H_1\bar{x}_0=0,\quad x_0'H_2\bar{x}_0=0$$

这又和条件 $x_0=0$ 矛盾．因此证明了 $q=0$ 或 $q=n-r$．即

$$H_1=\operatorname{diag}(I^{(p)},-I^{(r-p)},0).$$

$$H_2=\begin{pmatrix} A^{(r)} & B \\ \bar{B}' & \pm I^{(n-r)} \end{pmatrix}.$$

作复相合

$$\overline{\begin{pmatrix} I^{(r)} & 0 \\ \pm\bar{B}' & I \end{pmatrix}}'H_1\begin{pmatrix} I^{(r)} & 0 \\ \mp\bar{B}' & I \end{pmatrix}=H_1,$$

$$\overline{\begin{pmatrix} I^{(r)} & 0 \\ \mp\overline{B}' & I \end{pmatrix}}' H_2 \begin{pmatrix} I^{(r)} & 0 \\ \mp\overline{B}' & I \end{pmatrix} = \begin{pmatrix} A_1^{(r)} & 0 \\ 0 & \pm I \end{pmatrix}.$$

这样一来，问题化为将 $\text{diag}(I^{(p)}, \ -I^{(r-p)})$ 和 $A_1^{(r)}$ 在同一复相合下化为对角形. 换句话说，问题化为 $r=n$ 的情形，即在复相合意义下无妨设 n 阶 Hermite 方阵 H_1 非异.

设 H_1 和 H_2 为 n 阶 Hermite 方阵，且 $\det H_1 \neq 0$. 考虑多项式

$$\det(\lambda H_1 - H_2) = \det H_1 \det(\lambda I - H_1^{-1} H_2).$$

任取一根 λ_0，则存在 $0 \neq \alpha_n \in C^n$，使得

$$\lambda_0 H_1 \alpha_n = H_2 \alpha_n.$$

今由 $\det H_1 \neq 0$ 可知 $H_1 \alpha_n = \beta_n \neq 0$. 且无妨取 α_n，使得 β_n 为单位向量. 于是在 C^n 中存在标准正交基 $\beta_1, \ \beta_2, \ \cdots, \ \beta_n$. 由 $\overline{\beta}_j' \beta_n = \overline{\beta}_j' H_1 \alpha_n = 0$，所以 $\overline{\beta}_j' H_2 \alpha_n = 0, j=1,2,\cdots,n-1$. 这时 $\beta_1, \cdots, \beta_{n-1}$，$\alpha_n$ 线性无关. 事实上，若线性相关，则有 $\alpha_n = \sum_{j=1}^{n-1} a_j \beta_j$. 于是

$$\overline{\alpha}_n' H_i \alpha_n = \sum_{j=1}^{n-1} \overline{a}_j \overline{\beta}_j' H_i \alpha_n = 0, \quad i=1,2.$$

由条件可知 $\alpha_n = 0$，这和 $\alpha_n \neq 0$ 矛盾. 所以

$$P_0 = (\beta_1 \cdots \beta_{n-1} \alpha_n)$$

为 n 阶非异复方阵. 而

$$\overline{P}_0' H_i P_0 = \begin{pmatrix} \overline{\beta}_1' H_i \beta_1 & \cdots & \overline{\beta}_1' H_i \beta_{n-1} & \overline{\beta}_1' H_i \alpha_n \\ \vdots & & \vdots & \vdots \\ \overline{\beta}_{n-1}' H_i \beta_1 & \cdots & \overline{\beta}_{n-1}' H_i \beta_{n-1} & \overline{\beta}_{n-1}' H_i \alpha_n \\ \overline{\alpha}_n' H_i \beta_1 & \cdots & \overline{\alpha}_n' H_i \beta_{n-1} & \overline{\alpha}_n' H_i \alpha_n \end{pmatrix}$$

由 $\overline{\beta}_j' H_i \alpha_n = 0, i=1,2, j=1,2,\cdots n-1$，所以证明了

$$\overline{P}_0' H_i P_0 = \begin{pmatrix} \overline{H}_i^{(n-1)} & 0 \\ 0 & \overline{\alpha}_n' H_i \alpha_n \end{pmatrix}, \quad i=1,2.$$

今 $\det H_1 \neq 0$，所以$\det \widetilde{H}_1 \neq 0$. 由归纳法，便证明了存在 n 阶非异复方阵 Q_0，使得

$$\bar{Q}_0' H_1 Q_0 = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix}, \qquad \bar{Q}_0' H_2 Q_0 = \begin{pmatrix} \nu_1 & & \\ & \ddots & \\ & & \nu_n \end{pmatrix}.$$

由于 $\bar{Q}_0' H_i Q_0$ 是 Hermite 方阵，$i=1,2$，所以 $\mu_1, \cdots, \mu_n, \nu_1, \cdots, \nu_n$ 为实数. 由 $\det H_1 \neq 0$，再用一个置换方阵作复相合，从而无妨设 $\mu_1 \geqslant \cdots \geqslant \mu_p > 0 > \mu_{p+1} \geqslant \cdots \geqslant \mu_n$. 再用

$$\operatorname{diag}\left(\frac{1}{\sqrt{\mu_1}}, \cdots, \frac{1}{\sqrt{\mu_p}}, \frac{1}{\sqrt{-\mu_{p+1}}}, \cdots, \frac{1}{\sqrt{-\mu_n}}\right)$$

作复相合，便证明了当 H_1 非异时，Hermite 方阵偶 H_1, H_2在同一复相合下的标准形为

$$\operatorname{diag}(I^{(p)}, -I^{(n-p)}), \quad \operatorname{diag}(\lambda_1, \cdots, \lambda_n).$$

至此完全证明了定理.

定理 11.3.5 设$h(\alpha,\beta)$为 n 维线性空间 $\mathfrak{L}$ 上非退化Hermite 双线性函数. 则在 $\mathfrak{L}$ 中存在基$\{\beta_1, \beta_2, \cdots, \beta_n\}$，使得

$$\begin{aligned} h(\alpha, \beta) &= h\left(\sum_{j=1}^{n} x_j \beta_j, \ \sum_{k=1}^{n} y_k \beta_k\right) \\ &= \sum_{j=1}^{p} (1+\sqrt{-1}\lambda_j) x_j \bar{y}_j + \sum_{j=p+1}^{r} (-1+\sqrt{-1}\lambda_j) x_j \bar{y}_j \\ &\quad \pm \sqrt{-1} \sum_{j=r+1}^{n} x_j \bar{y}_j. \end{aligned}$$

证 今在 $\mathfrak{L}$ 中任取一组基 $\alpha_1, \cdots, \alpha_n$，则

$$h(\alpha, \beta) = h\left(\sum_{j=1}^{n} \mu_j \alpha_j, \ \sum_{k=1}^{n} \nu_k \alpha_k\right) = \mu'(H_1 + \sqrt{-1} H_2)\bar{\nu}$$

由非退化条件，有 $\mu' H_1 \bar{\mu} = 0, \mu' H_2 \bar{\mu} = 0$，则 $\mu = 0$. 由定理11.3.4，存在基$\{\beta_1, \cdots, \beta_n\}$，使得

$$h(\alpha,\beta)=h\left(\sum_{j=1}^{n}x_j\beta_j,\ \sum_{k=1}^{n}y_k\beta_k\right)$$

$$=x'\left[\begin{pmatrix}I^{(p)} & 0 & 0\\ 0 & -I^{(r-p)} & 0\\ 0 & 0 & 0\end{pmatrix}+\sqrt{-1}\begin{pmatrix}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{pmatrix} & 0\\ 0 & \pm I\end{pmatrix}\right]\bar{y}.$$

$$=x'\mathrm{diag}((1+\sqrt{-1}\lambda_1,\cdots,1+\sqrt{-1}\lambda_p),(-1+\sqrt{-1}\lambda_{p+1},$$
$$\cdots,-1+\sqrt{-1}\lambda_r),\ \pm\sqrt{-1}I)\bar{y}$$

$$=\sum_{j=1}^{p}(1+\sqrt{-1}\lambda_j)x_j\bar{y}_j+\sum_{j=p+1}^{r}(-1+\sqrt{-1}\lambda_j)x_j\bar{y}_j$$
$$\pm\sqrt{-1}\sum_{j=r+1}^{n}x_j\bar{y}_j.$$

下面在 n 维复线性空间 Ω 上考虑复对称双线性函数 $f(\alpha,\beta)$. 于是有

定义 设 $f(\alpha,\beta)$ 为 n 维复线性空间 Ω 上对称双线性函数,则 $f(\alpha,\alpha)$, $\forall\alpha\in\Omega$ 称为**复二次型**.

显然,在 n 维复线性空间 Ω 中取定一组基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 记 $\alpha=\sum_{j=1}^{n}x_j\alpha_j$, 则复二次型

$$f(\alpha,\alpha)=x'Sx,$$

其中 S 为复对称方阵:

$$S=\begin{pmatrix}f(\alpha_1,\alpha_1) & \cdots & f(\alpha_1,\alpha_n)\\ \vdots & & \vdots\\ f(\alpha_n,\alpha_1) & \cdots & f(\alpha_n,\alpha_n)\end{pmatrix},\quad x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}.$$

定义 n 阶复方阵 A 和 B 称为**相合的**, 如果存在 n 阶非异复方阵 P, 使得

$$B=P'AP$$

引理 11.3.4 相合为等价关系.

证明由直接验证可知.

显然,复对称(复斜对称)方阵在相合下仍变为复对称(复斜对称)方阵,且当方阵全为实时,相合即为实相合. 由此可知,从实的情形推广为复的情形,出现两种类型: 一是复相合,一是相合. 从结果来看,复相合是实相合的自然推广,而相合并不是实相合的自然推广,事实上,我们有

定理 11.3.6 n 阶复对称方阵 S 必相合于标准形

$$\operatorname{diag}(I^{(r)},0)$$

其中 $r=\operatorname{rank}(S)$. 所以 n 阶复对称方阵的秩为它在相合下的全系不变量.

证 用定理 10.1.1 的第二个证明可知,存在 n 阶非异复方阵 P_1,使得

$$P_1'SP_1=\operatorname{diag}(I^{(p)},-I^{(r-p)},0).$$

取

$$P=P_1\operatorname{diag}(I^{(p)},\sqrt{-1}I^{(r-p)},I^{(n-r)}),$$

则有

$$P'SP=\operatorname{diag}(I^{(r)},0).$$

证完.

作为推论,我们有

定理 11.3.7 设 $\varphi(\alpha)$ 为 n 维复线性空间 Ω 上复二次型,则在 Ω 中存在一组基 $\beta_1,\beta_2,\cdots,\beta_n$,使得

$$\varphi(\alpha)=\varphi\left(\sum_{j=1}^{n}x_j\beta_j\right)=\sum_{j=1}^{r}x_j^2.$$

和定理 10.3.1 一样可证

定理 11.3.8 对 n 阶复斜对称方阵 K,则存在 n 阶非异复方

阵P,使得

$$P'KP=\operatorname{diag}\left(\begin{pmatrix}0 & a_1\\ -a_1 & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & a_s\\ -a_s & 0\end{pmatrix},0,\cdots,0\right)$$

其中 $a_1,\cdots,a_s$ 为非零复数，且 $\det P=1$. 又 P 及 $P'KP$ 中元素都是 K 中元素的有理分式，因此 n 阶复斜对称方阵 K 相合于标准形

$$\operatorname{diag}\left(\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},0,\cdots,0\right).$$

特别复斜对称方阵的秩必为偶数，且它是 n 阶复斜对称方阵在相合下的全系不变量.

习题*11.3

1. 试将下列 Hermite 方阵在复相合下化为标准形.

(i) $\begin{pmatrix}1 & 1+i & \cdots & (n-1)+i\\ 1-i & 1 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ (n-1)-i & 0 & \cdots & 1\end{pmatrix}$,

(ii) $\begin{pmatrix}1 & \frac{i}{2} & & & \\ -\frac{i}{2} & 1 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 1 & \frac{i}{2}\\ & & & -\frac{i}{2} & 1\end{pmatrix}$,

其中 $i=\sqrt{-1}$.

2. 试将下列 Hermite 型化为标准形.

(i) $\sum\limits_{j=1}^{n-1}(x_j+\bar{x}_{j+1}+\bar{x}_jx_{j+1})$,　　(ii) $\sum\limits_{j,k=1}^{n}|j-k|x_j\bar{x}_k$.

3. 试将下列复对称方阵在相合下化为标准形.

(i) $\begin{pmatrix} 1 & 1+i & \cdots & (n-1)+i \\ 1+i & 1 & \cdots & 0 \\ \vdots & \vdots & & 0 \\ (n-1)+i & 0 & \cdots & 1 \end{pmatrix}$,

(ii) $\begin{pmatrix} 1 & \frac{i}{2} & & & \\ \frac{i}{2} & 1 & \ddots & & \\ & & \ddots & & \\ & & \ddots & 1 & \frac{i}{2} \\ & & & \frac{i}{2} & 1 \end{pmatrix}$,

其中 $i=\sqrt{-1}$.

§11.4 定正 Hermite 方阵和方阵的极分解

显然 Hermite 方阵的定正(半定正、定负、半定负)的性质在复相合下不改变.

定理 11.4.1 设 H 为 n 阶 Hermite 方阵，则下面性质相互等价.

(1) $H>0$;

(2) H 的特征根都是正实数;

(3) 存在非异复方阵 Q,使得 $H=\bar{Q}'Q$;

(4) H 的 n 个顺序主子式 $H\binom{12\cdots j}{12\cdots j}>0, j=1,2,\cdots,n$;

(5) H 的一切主子式都是正实数.

定理 11.4.2 设 H 为 n 阶 Hermite 方阵，则下面性质互相等价.

(1) $H\geqslant 0$;

(2) H 的特征根都是非负实数;

(3) 存在 n 阶复方阵 Q，使得 $H=\bar{Q}'Q$；

(4) H 的一切主子式都是非负实数.

定理 11.4.3 设 H 为 n 阶半定正Hermite 方阵，记 $r=\text{rank}(H)$，则存在 r 个线性无关的 $n\times 1$ 复矩阵 $\beta_1,\beta_2,\cdots,\beta_r$，使得

$$H=\sum_{j=1}^{n}\beta_j\bar{\beta}_j'.$$

定理 11.4.4 设 H 为 n 阶半定正 Hermite 方阵，则唯一存在 n 阶半定正 Hermite 方阵 H_1，使得

$$H_1^2=H.$$

且任一 n 阶复方阵 A，若有 $AH=HA$，则必有 $AH_1=H_1A$. 再 $H>0$，当且仅当 $H_1>0$.

定理 11.4.5(极分解式) 设 A 为 n 阶复方阵，则存在 n 阶酉方阵 U 和 n 阶半定正 Hermite 方阵 H_1 和 H_2，使得

$$A=H_1U=UH_2.$$

这称为方阵 A 的**极分解**. 当 A 非异时，则 U,H_1,H_2 唯一存在.

定义 设 A 和 B 为 $n\times m$ 复矩阵，若存在 n 阶酉方阵 U 和 m 阶酉方阵 V，使得 $B=UAV$，则称 A 和 B 为**酉相抵的**.

引理 11.4.1 $n\times m$ 复矩阵的酉相抵为等价关系.

定理 11.4.6 设 A 为 $n\times m$ 复矩阵，则存在 n 阶酉方阵 U 和 m 阶酉方阵 V，使得

$$UAV=\begin{pmatrix}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_r\end{pmatrix} & 0\\ 0 & 0\end{pmatrix},$$

其中 $\lambda_1\geqslant\cdots\geqslant\lambda_r>0$，$r=\text{rank}(A)$，又 $\lambda_1^2,\cdots,\lambda_r^2$ 为 $A\bar{A}'$的所有非零特征根. 所以 $A\bar{A}'$的所有非零特征根为 A 在酉相抵的全系不变量.

推论 设 A 为 $n\times m$ 复矩阵，则

$$\text{rank}(A)=\text{rank}(A\bar{A}')=\text{rank}(\bar{A}'A).$$

且 $A\bar{A}'$ 和 $\bar{A}'A$ 的非零特征根相同.

上面一系列结论的证明都和 §10.2 相同.

习题 11.4

1. 习题 10.2 中哪些结论可以推广到复的情形.

2. 设 H 为半定正 Hermite 方阵，记 A,B 分别为它的实部及虚部，则 A 实对称，B 实斜对称. 又

$$\max(\operatorname{rank}(H),\operatorname{rank}(B))\leqslant\operatorname{rank}(A).$$

3. 设 H 为 n 阶定正 Hermite 方阵，试证：

$$\det H\leqslant H\binom{1}{1}H\binom{2\ \cdots\ n}{2\ \cdots\ n}\leqslant\cdots\leqslant H\binom{1}{1}H\binom{2}{2}\cdots H\binom{n}{n}.$$

且有

$$\det H\leqslant\det\operatorname{Re}H.$$

其中等号成立当且仅当 H 为实定正对称方阵. 且有

$$\det\operatorname{Im}H<\det\operatorname{Re}H.$$

4. 设 H 为 $2n$ 阶定正 Hermite 方阵，且 $HJ=J\bar{H}$，其中

$$J=\begin{pmatrix}0 & I^{(n)}\\ -I^{(n)} & 0\end{pmatrix}.$$

试证：存在 n 阶非异复方阵 Q，使得 $QJ=J\bar{Q}$，$H=\bar{Q}'Q$. 再证 Q 可取非异上三角方阵.

5. 设 A 为 n 阶复方阵，记作 $A=(a_{ij})$. 试证：

$$|\det A|^2\leqslant\prod_{j=1}^{n}\sum_{k=1}^{n}|a_{jk}|^2.$$

6. 设 H_1,H_2 为一对 Hermite 方阵，其中 $H_1>0$. 试证：存在非异复方阵 P，使得

$$H_1=\bar{P}'P,\qquad H_2=\bar{P}'\operatorname{diag}(\lambda_1,\cdots,\lambda_n)P,$$

其中 $\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n$.

7. 设 H_1,H_2 都是定正 Hermite 方阵. 试证：

(i) 当 $H_1-H_2>0$，则有 $H_2^{-1}-H_1^{-1}>0$;

(ii) $(\det H_1)^{\frac{1}{n}}+(\det H_2)^{\frac{1}{n}}\leqslant(\det(H_1+H_2))^{\frac{1}{n}}$.

又等号成立当且仅当 $H_2=aH_1$，其中 a 为正实数.

8. 设 A 为 $n\times m$ 复矩阵，B 为 $m\times p$ 复矩阵，且 $n=\operatorname{rank}(A)$. 试证：

$$\det[\bar{B}'(I-\bar{A}'(A\bar{A}')^{-1}A)B]\leqslant\det\bar{B}'B.$$

9. 设 A 和 B 为 n 阶复方阵,使得 $I-A\bar{A}'\geqslant 0, I-B\bar{B}'\geqslant 0$,则有

$$|\det(I-A\bar{B}')|^2\geqslant|\det(I-A\bar{A}')\det(I-B\bar{B}')|.$$

10. 设 A 和 B 都是 Hermite 方阵,且 A 定正. 试证: $\det(\lambda A+B)$ 的根都是实数.

11. 设 A 和 B 都是 n 阶复对称方阵, $\det B\neq 0$. 设多项式 $\det(A+\lambda B)$ 没有重根. 试证: A 和 B 可以同时相合于对角形.

12. 设 A 和 B' 为 $n\times m$ 复矩阵. 试证: AB 和 BA 同时为规范方阵当且仅当 $\bar{A}'AB=BA\bar{A}', \bar{B}'BA=AB\bar{B}'$.

§ 11.5 复方阵在酉相合下的标准形

定义 两个 n 阶复方阵 A 和 B 称为**酉相合的**,如果存在 n 阶酉方阵 U,使得

$$B=\bar{U}'AU.$$

引理 11.5.1 n 阶复方阵的酉相合关系为等价关系.

引理 11.5.2 给定 n 阶复方阵 A 及 n 阶酉方阵 U. 设 $UA=AU'$,则存在酉方阵 V,使得 $VA=AV'$,且 $V^2=U$.

证 今存在 n 阶酉方阵 U_1 使得

$$\bar{U}_1'UU_1=\mathrm{diag}(e^{\sqrt{-1}\theta_1},\cdots,e^{\sqrt{-1}\theta_n}).$$

记

$$V=U_1\mathrm{diag}(e^{\frac{\sqrt{-1}\theta_1}{2}},\cdots,e^{\frac{\sqrt{-1}\theta_n}{2}})\bar{U}_1'.$$

于是有 $V^2=U$,且 V 为酉方阵,余下要证 $VA=AV'$.

今已知 $UA=AU'$,即有

$$U_1\mathrm{diag}(e^{\sqrt{-1}\theta_1},\cdots,e^{\sqrt{-1}\theta_n})\bar{U}_1'A=A\bar{U}_1\mathrm{diag}(e^{\sqrt{-1}\theta_1},\cdots,e^{\sqrt{-1}\theta_n})U_1'$$

于是有

$$\mathrm{diag}(e^{\sqrt{-1}\theta_1},\cdots,e^{\sqrt{-1}\theta_n})(\bar{U}_1'A\bar{U}_1)=(\bar{U}_1'A\bar{U}_1)\mathrm{diag}(e^{\sqrt{-1}\theta_1},\cdots,e^{\sqrt{-1}\theta_n}).$$

记$\bar{U}'_1 A\bar{U}_1=(b_{jk})$，此即

$$(e^{\sqrt{-1}\theta_i}-e^{\sqrt{-1}\theta_j})b_{ij}=0.$$

当$b_{ij}\neq 0$时有$e^{\sqrt{-1}\theta_i}=e^{\sqrt{-1}\theta_j}$，因此$e^{\frac{\sqrt{-1}\theta_i}{2}}=e^{\frac{\sqrt{-1}\theta_j}{2}}$. 所以证明了

$$(e^{\frac{\sqrt{-1}\theta_i}{2}}-e^{\frac{\sqrt{-1}\theta_j}{2}})b_{ij}=0,\quad 1\leqslant i,j\leqslant n.$$

此即

$$\operatorname{diag}(e^{\frac{\sqrt{-1}\theta_1}{2}},\cdots,e^{\frac{\sqrt{-1}\theta_n}{2}})\bar{U}'_1 A\bar{U}_1$$

$$=\bar{U}'_1 A\bar{U}_1\operatorname{diag}(e^{\frac{\sqrt{-1}\theta_1}{2}},\cdots,e^{\frac{\sqrt{-1}\theta_n}{2}}).$$

由V之定义可知$VA=AV'$. 证完.

定理 11.5.1 n阶复对称方阵S必酉相合于对角形

$$\Lambda=\operatorname{diag}(\lambda_1,\cdots,\lambda_n),$$

其中$\lambda_1\geqslant\cdots\geqslant\lambda_n\geqslant 0$，又$\lambda_1^2,\cdots,\lambda_n^2$为$n$阶 Hermite 方阵$S\bar{S}'$的所有特征根. 所以$S\bar{S}'$的所有特征根为$S$在酉相合下的全系不变量.

证 由定理 11.4.5 存在n阶酉方阵U_1,V_1，使得

$$U_1SV_1=\Lambda.$$

由$\Lambda'=\Lambda$，所以$U_1SV_1=(U_1SV_1)'=V'_1SU'_1$，即有

$$\bar{V}_1U_1S=SU'_1\bar{V}'_1=S(\bar{V}_1U_1)'.$$

由引理 11.5.2，存在n阶酉方阵W，使得

$$WS=SW',\qquad W^2=\bar{V}_1U_1.$$

取n阶酉方阵

$$U=W'V_1,$$

则有

$$U'SU=V_1'WSW'V_1=V_1'W^2SV_1$$
$$=V_1'\bar{V}_1U_1SV_1=U_1SV_1=\Lambda.$$

这证明了定理. 证完.

引理 11.5.3 设 T 为 n 阶非异复斜对称方阵,则 $T\bar{T}'$ 的特征根必为重根.

证 今 $\det T\neq 0, -T'=T$. 所以 T 为偶数阶的复斜对称方阵,由

$$\det(\lambda I-T\bar{T}')=\det(\lambda T^{-1}-\bar{T}')\det T.$$

对复斜对称方阵 $\lambda T^{-1}-\bar{T}'$ 用定理 11.3.8,所以存在行列式等于 1 的 n 阶复方阵 P,使得

$$P'(\lambda T^{-1}-\bar{T}')P=\mathrm{diag}\left(\begin{pmatrix}0 & a_1\\ -a_1 & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & a_r\\ -a_r & 0\end{pmatrix}\right),$$

其中 $r=\frac{n}{2}$,且 $a_1,\cdots, a_r$ 及 P 中所有元素都是 $\lambda T^{-1}-\bar{T}'$ 中元素的有理分式,即为 λ 的有理分式,所以

$$a_1a_2\cdots a_r=\frac{f(\lambda)}{g(\lambda)},$$

其中 $f(\lambda),g(\lambda)$ 为互素的多项式,且 $g(\lambda)\neq 0$ 的首项系数为 1. 而

$$(\det P)^2\det(\lambda T^{-1}-\bar{T}')=a_1^2a_2^2\cdots a_r^2=\frac{f(\lambda)^2}{g(\lambda)^2}.$$

由 $\det P=1$, $\det(\lambda T^{-1}-\bar{T}')$ 为 λ 的 n 次多项式可知 $g^2|f^2$. 由 $(f,g)=1$ 可知 $g=1$. 这也证明了

$$\det(\lambda I-T\bar{T}')=\det T\det(\lambda T^{-1}-\bar{T}')$$
$$=(\det T)f(\lambda)^2.$$

由 $\det T$ 为复常数,所以 $\det(\lambda I-T\bar{T}')$ 为复系数多项式的平方. 因此 $\det(\lambda I-T\bar{T}')$ 的根都是重根. 证完.

定理 11.5.2 复斜对称方阵 T 必酉相合于标准形

$$\operatorname{diag}\left(\begin{pmatrix} 0 & a_1 \\ -a_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & a_s \\ -a_s & 0 \end{pmatrix}, 0, \cdots, 0\right),$$

其中 $0<a_s\leqslant\cdots\leqslant a_1$，又 $a_1^2, a_1^2, \cdots, a_s^2$，$a_s^2$ 为 TT' 的所有非零特征根. 它们是 T 在酉相合下的全系不变量.

证 考虑齐次线性方程组 $Tx=0$，它的解空间为 C^n 中的 $n-\operatorname{rank}(T)$ 维子空间，在其中任取标准正交基 $\beta_{r+1}, \beta_{r+2}, \cdots, \beta_n$，其中 $r=\operatorname{rank}(T)$. 于是在 C^n 中存在标准正交基 $\beta_1, \beta_2, \cdots, \beta_r, \beta_{r+1}, \cdots, \beta_n$. 记

$$U=(\beta_1\beta_2\cdots\beta_r\beta_{r+1}\cdots\beta_n).$$

则 U 为 n 阶酉方阵. 且由 $T\beta_{r+1}=\cdots=T\beta_n=0$，所以有

$$TU=(T_1^{(n,r)}, 0)$$

因此由 $U'TU$ 为 n 阶斜对称方阵，所以有

$$U'TU=\begin{pmatrix} T_0^{(r)} & 0 \\ 0 & 0 \end{pmatrix}.$$

其中 T_0 为 r 阶非异复斜对称方阵.

因此问题化为 T 为非异的情形来证明定理. 由引理 11.5.3，所以 TT' 的特征根为

$$a_1^2=a_1^2\geqslant a_2^2=a_2^2\geqslant\cdots\geqslant a_s^2=a_s^2,$$

其中 $s=\frac{n}{2}$，$a_1\geqslant a_2\geqslant\cdots\geqslant a_s>0$，且存在 n 阶酉方阵 U_1 和 U_2，使得

$$U_1TU_2=\operatorname{diag}(a_1, a_1, \cdots, a_s, a_s).$$

取 n 阶酉方阵 $U_3=\operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)$，则有

$$U_3U_1TU_2=\operatorname{diag}\left(\begin{pmatrix} 0 & a_1 \\ -a_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & a_s \\ -a_s & 0 \end{pmatrix}\right).$$

因此 $U_3U_1TU_2+U_2'T'U_1'U_3'=0$，此即 $\bar{U}_2U_3U_1T=T(\bar{U}_2U_3U_1)'$.由引理 11.5.2,存在 n 阶酉方阵 V,使得

$$V^2=\bar{U}_2U_3U_1, \qquad VT=TV'.$$

令

$$U=V'U_2,$$

则有

$$\begin{aligned} U'TU &= U_2'VTV'U_2=U_2'V^2TU_2=U_3U_1TU_2 \\ &= \mathrm{diag}\left(\begin{pmatrix} 0 & a_1 \\ -a_1 & 0 \end{pmatrix},\cdots,\begin{pmatrix} 0 & a_s \\ -a_s & 0 \end{pmatrix}\right). \end{aligned}$$

这证明了 $U'TU$ 为标准形. 定理证完.

第十二章　广义逆矩阵

§ 12.1　强广义逆矩阵

为了引进强广义逆矩阵．先建立 n 维点空间及其上的解析几何学．

熟知平面解析几何学是研究如何将平面中几何性质化为代数问题，且解决此代数问题．办法是在普通平面中引进直角坐标系，将平面中点与实数偶(即此点之坐标)间建立一一对应．对平面中任两点 $p=(a_1,a_2)$，$q=(b_1,b_2)$，以 p，q 为端点之线段记作 $\overline{pq}$，以 p 为起点，q 为终点之向量记作 $\overrightarrow{pq}$．于是线段 $\overline{pq}$ 之长度等于向量 $\overrightarrow{pq}$ 之长度，定义为

$$d(p,q)=|\overline{pq}|=|\overrightarrow{pq}|=\sqrt{(a_1-b_1)^2+(a_2-b_2)^2}.$$

向量 $\overrightarrow{pq}$ 之坐标为 (b_1-a_1,b_2-a_2)．平面上两向量的坐标相等，则称为向量相等．将所有相等的向量看作一个向量，则此向量称为自由向量．确切地说，向量相等为等价关系，所有平面上向量按此等价关系分成了等价类．每一类可用其中之向量之坐标来表示，称为自由向量．虽然，平面上所有自由向量构成一个 2 维实线性空间．它的任一个子空间中向量，都安在平面上一个固定点 p 上，即都以 p 点为起点，则终点构成平面中子集合，称为子(点)空间．平面称为 2 维实(点)空间，或 2 维 Euclid(点)空间．

同理，空间解析几何学是研究如何将空间中几何性质化为代数问题，且解决此代数问题．办法是在普通空间中引进直角坐标

系，将空间中点与三实数(即此点之坐标)间建立了一一对应. 向量，自由向量，线段及其长度之定义和平面情形一样，这也就启发我们可以抽象地引进 n 维实(复)点空间，它也称为 n 维Euclid(复Euclid) 点空间，复 Euclid 点空间也可称为 n 维酉点空间. 由于和普通情形不一样，没有一个几何实体，我们就用点坐标来代替.

定义 所有 $n\times 1$ 实(复) 矩阵构成之集合称为 n **维实(复)点空间**，记作 $R^n(C^n)$. 其中元素称为点，若点

$$p=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},$$

则 $\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$ 称为点 p 之坐标，a_i 称为第 i 个坐标分量. 在点空间中任取两点

$$p=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},\qquad q=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix},$$

则非负整数

$$d(p,q)=\sqrt{\sum_{i=1}^{n}(b_i-a_i)^2}$$

称为点 p 与点 q 间之**距离**，或以点 p，q 为端点之线段 $\overline{pq}$ 之**长度**，以点 p 为起点，点 q 为终点的向量 $\overrightarrow{pq}$ 之坐标定义为 $n\times 1$ 矩阵

$$\boldsymbol{\alpha}=\begin{pmatrix}b_1-a_1\\b_2-a_2\\\vdots\\b_n-a_n\end{pmatrix}.$$

点空间中坐标相同之向量称为相等的向量．所有相等的向量可以用同一个坐标来表示，抽象地把所有相等的向量看作一个几何对象，则称为**自由向量**．点空间中所有自由向量(仍用 $n\times 1$ 实(复)矩阵来表示)构成 n 维实(复) 线性空间 V_n．按第九章和第十一章的办法在其中引进标准内积，则 V_n 为实(复)Euclid 空间．

任取 V_n 中子空间 $\mathfrak{L}$，将 $\mathfrak{L}$ 中自由向量全都安在点空间 $R^n(C^n)$ 中一个固定点 p 上，则终点全体构成之集合称为包含点 p 之子(点)空间，记作 $(\mathfrak{L},p)$．显然，子(点)空间仍为点空间，其维数等于 V_n 之线性子空间 $\mathfrak{L}$ 之维数．当 p 取作原点(即令 $p=0$)时，子(点)空间 $(\mathfrak{L},0)$ 中点 q 之坐标和向量 $\overrightarrow{0q}$ 之坐标相同．

点空间 $R^n(C^n)$ 中任取两线段 $\overline{pq}$ 和 $\overline{rs}$，它们称为**互相正交的**，如果自由向量 $\overrightarrow{pq}$ 和 $\overrightarrow{rs}$ 在 V_n 中互相正交；称为互相平行的，如果自由向量 $\overrightarrow{pq}$ 和 $\overrightarrow{rs}$ 在 V_n 中相等．

引理 12.1.1 点空间中任取三点 p,q,r，则距离有性质

(1) $d(p,q)\geqslant 0$，且等号成立当且仅当 $p=q$；

(2) $d(p,q)=d(q,p)$；

(3) $d(p,r)\leqslant d(p,q)+d(q,r)$ (三角不等式)．

证 由距离的定义可知性质(1)，(2)成立．下面来证性质(3)．记自由向量 $\overrightarrow{0p}=\alpha,\overrightarrow{0q}=\beta,\ \overrightarrow{0r}=\gamma$，由距离之定义可知 $d(p,\ q)=|\overrightarrow{0p}-\overrightarrow{0q}|=|\alpha-\beta|$．于是问题化为证

$$|\alpha-\gamma|\leqslant|\alpha-\beta|+|\beta-\gamma|,\quad \forall\alpha,\beta,\gamma\in V_n.$$

今由 Cauchy 不等式，

$$\begin{aligned}|\alpha-\gamma|^2&=((\alpha-\beta)+(\beta-\gamma),(\alpha-\beta)+(\beta-\gamma))\\&=|\alpha-\beta|^2+|\beta-\gamma|^2+2\mathrm{Re}(\alpha-\beta,\beta-\gamma)\\&\leqslant|\alpha-\beta|^2+|\beta-\gamma|^2+2|(\alpha-\beta,\beta-\gamma)|\\&\leqslant|\alpha-\beta|^2+|\beta-\gamma|^2+2|\alpha-\beta||\beta-\gamma|\end{aligned}$$

$$=(|\alpha-\beta|+|\beta-\gamma|)^2.$$

证完.

定义 在 n 维点空间 $R^n(C^n)$ 中给定子(点)空间 $(\mathfrak{L},p)$ 及一定点 q. 则点 q 到子(点)空间 $(\mathfrak{L},p)$ 之距离为

$$\min_{r\in(\mathfrak{L},p)} d(q,r).$$

关于点空间中定点到定子(点)空间之距离有下面性质

引理 12.1.2 在 n 维点空间 $R^n(C^n)$ 中给定一点 q 及一子(点)空间 $(\mathfrak{L},p)$. 记 $\mathfrak{L}^\perp=\{\xi\in V_n|(\xi,\mathfrak{L})=0\}$ 为 $\mathfrak{L}$ 之正交补，则有 $V_n=\mathfrak{L}\oplus\mathfrak{L}^\perp$，特别 $\overrightarrow{pq}$ 可唯一地分解为:

$$\overrightarrow{qp}=\alpha+\beta,$$

其中 $\alpha\in\mathfrak{L}^\perp,\beta\in\mathfrak{L}$. 于是在子(点)空间 $(\mathfrak{L},p)$ 中唯一存在一点 s，使得 $\overrightarrow{qs}=\alpha$, 且

$$\min_{r\in(\mathfrak{L},p)} d(q,r)=|\overrightarrow{qs}|=|\overline{qs}|.$$

证 在子(点)空间 $(\mathfrak{L},p)$ 中任取一点 r, 则 $\overrightarrow{pr}\in\mathfrak{L}$. 于是

$$\overrightarrow{qr}=\overrightarrow{qp}+\overrightarrow{pr}=\alpha+(\beta+\overrightarrow{pr}),$$

其中 $\alpha\in\mathfrak{L}^\perp$, $\beta+\overrightarrow{pr}\in\mathfrak{L}$, 于是

$$d(q,r)^2=|\overrightarrow{qr}|^2=(\alpha+(\beta+\overrightarrow{pr}),\ \alpha+(\beta+\overrightarrow{pr}))$$
$$=|\alpha|^2+|\beta+\overrightarrow{pr}|^2\geqslant|\alpha|^2,\ \forall r\in(\mathfrak{L},p).$$

特别，由 $-\beta\in\mathfrak{L}$，所以唯一存在 $s\in(\mathfrak{L},p)$ 使得 $\overrightarrow{ps}=-\beta$，因此 $d(q,s)^2=|\alpha|^2+|\beta+\overrightarrow{ps}|^2=|\alpha|^2$. 这证明了

$$\min_{r\in(\mathfrak{L},p)} d(q,r)=|\alpha|=d(q,s)=|\overrightarrow{qs}|.$$

今 $\alpha=\overrightarrow{qp}-\beta=\overrightarrow{qp}+\overrightarrow{ps}=\overrightarrow{qs}$. 引理证完.

关于 n 维点空间中超平面，二次超曲面等等定义及性质，读者

可以参照普通解析几何学的办法形式推广，也可参考拙著《代数学引论》一书.

下面先介绍最小二乘解，然后引进强广义逆矩阵.

由于在实际观测时，不可避免地要出现各种误差，例如人为观察造成的误差，仪器产生之误差，以及用四舍五入法处理数据引起的误差等等. 所以由观测数据导出的线性方程组

$$\sum_{k=1}^{m} a_{jk}x_k = b_j,\ \ 1\leqslant j\leqslant n$$

中 a_{jk} 及 b_j 都是近似值. 因此或者出现线性方程组不相容的情形，或者虽然线性方程组相容，但是得到之解并不是能反映观测对象的客观实际. 为了尽可能准确地去反映客观实际，我们可以寻找某种意义下误差最小的解. 用统计学中数据处理的办法，最自然的是最小二乘解.

记

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix},\quad x=\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix},\quad \beta=\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

于是 $n\times m$ 矩阵 A 给出了 V_m 到 V_n 内的映射，它定义为 $\mathscr{A}: x\rightarrow Ax$，$\forall x\in V_m$. 今 V_n 中子集合 $\mathscr{A}(V_m)$ 定义了 V_m 中过原点 0 之子空间 $(\mathscr{A}(V_m), 0)$，β 为 C^n 中一定点. 由引理 12.1.2，在 C^n 之子空间 $(\mathscr{A}(V_m), 0)$ 中唯一存在一点 $s=\mathscr{A}(x_0)=Ax_0$，其中

$$x_0=\begin{pmatrix} x_1^{(0)} \\ \vdots \\ x_m^{(0)} \end{pmatrix},$$

使得 $d(s,\beta)$ 为点 β 到子(点)空间 $(\mathscr{A}(V_m), 0)$ 之距离，此即

$$|Ax_0-\beta|^2\leqslant|Ax-\beta|^2,\quad \forall x\in C^m.$$

所以 x_0 称为线性方程组 $Ax_0=\beta$ 之**最小二乘解**. 显然，最小二乘

解不唯一，其中解 x_0 之模 $|x_0|$ 最小之解为**模最小解**.

下面来求最小二乘解和模最小解.

引理 12.1.3 设 A 为 $n\times m$ 矩阵，$\beta\in V_n$. $x_0\in V_m$ 为线性方程组

$$Ax=\beta$$

的最小二乘解当且仅当 x_0 为相容线性方程组

$$\bar{A}'Ax=\bar{A}'\beta$$

的解.

证 先证线性方程组 $\bar{A}'Ax=\bar{A}'\beta$ 相容. 事实上，由于 $\mathrm{rank}(\bar{A}'A)=\mathrm{rank}(A)$，而

$$\begin{aligned}\mathrm{rank}(A)=\mathrm{rank}(\bar{A}'A)&\leqslant\mathrm{rank}((\bar{A}'A\bar{A}'\beta))\\&=\mathrm{rank}(\bar{A}'(A\beta))\\&\leqslant\min(\mathrm{rank}(\bar{A}'),\mathrm{rank}((A\ \ \beta)))\\&=\min(\mathrm{rank}(A),\mathrm{rank}((A\ \ \beta)))\\&=\mathrm{rank}(A),\end{aligned}$$

这证明了 $\mathrm{rank}(\bar{A}'A)=\mathrm{rank}((\bar{A}'A\ \ \bar{A}'\beta))$，即线性方程组 $\bar{A}'Ax=\bar{A}'\beta$ 相容.

下面来求线性方程组 $Ax=\beta$ 的最小二乘解. 在 V_m 及 V_n 中取标准基，设点 q 坐标为 β，$n\times m$ 矩阵 A 决定的线性映射为 $\mathscr{A}:V_m\to V_n$. 由引理12.1.2，记 $\mathscr{A}(V_m')$ 在 V_n 中的正交外为 $(\mathscr{A}(V_m))^{\perp}$. 于是在点空间 C^n 的子点空间 $(\mathscr{A}(V_m),0)$ 中唯一存在一点 $s=Ax_0$，其中 x_0 在 V_m 中，使得 $\overrightarrow{q0}=\overrightarrow{qs}+\overrightarrow{s0}$. 而 $\overrightarrow{qs}\in(\mathscr{A}(V_m))^{\perp}$，$\overrightarrow{s0}\in\mathscr{A}(V_m)$. 且 $|\overrightarrow{qs}|$ 为 q 到 $(\mathscr{A}(V_m),0)$ 之距离.

今 $s=Ax_0$，于是 $\overrightarrow{qs}=Ax_0-\beta$；对 $r=Ax\in(\mathscr{A}(V_m),\ 0)$，$\forall x\in C^m$，所以 $\overrightarrow{qr}=Ax-\beta$. 由引理 12.1.2，有 $|\overrightarrow{qs}|\leqslant|\overrightarrow{qr})$， 即有 $|Ax_0-\beta|\leqslant|Ax-\beta|$，$\forall x\in C^m$，此即 x_0 为 $Ax=\beta$ 之最小二乘解.

今 $\overrightarrow{\beta s}\in(\mathscr{A}(V_m))^{\perp}$ 当且仅当 $(\overrightarrow{qs},\mathscr{A}(V_m))=0$，此即 $(Ax_0-\beta, Ax)=0$，$\forall x\in V_m$. 由于内积为标准内积，写成矩阵形式，即 $(\overline{Ax})'(Ax_0-\beta)=0$，此即 $\bar{x}'(\bar{A}'Ax_0-\bar{A}'\beta)=0$，$\forall x\in V_m$，所以 $\bar{A}'Ax_0=\bar{A}'\beta$. 这证明了 $x_0\in V_m$ 为 $Ax=\beta$ 之最小二乘解当且仅当 $x_0\in V_m$ 为 $\bar{A}'Ax=\bar{A}'\beta$ 之解. 引理证完.

定义 给定 $n\times m$ 复矩阵 A，记 $\bar{A}'A$ 的非零特征根为 $\lambda_1\geqslant\cdots\geqslant\lambda_r>0$. 记 n 阶酉方阵 U 及 m 阶酉方阵 V，使得

$$A=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V,\quad \Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_r),$$

则 $m\times n$ 复矩阵

$$A^{+}=V^{-1}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}U^{-1}$$

称为 $n\times m$ 复矩阵 A 的**强广义逆矩阵**.

显然，当 A 为 n 阶非异方阵时，强广义逆矩阵就是普通的逆矩阵. 从这个角度来说，强广义逆矩阵是把逆方阵的概念推广到奇异方阵及一般矩阵的情形. 另外，上面定义依赖于 U 及 V 之选取，但 U,V 并不唯一，这样定义的强广义逆矩阵是否确定呢？我们的回答是肯定的，即有

引理 12.1.4 设 U,U_1 为 n 阶酉方阵，V,V_1 为 m 阶酉方阵，且有

$$A=U\begin{pmatrix}\Lambda^{(r)} & 0\\ 0 & 0\end{pmatrix}V=U_1\begin{pmatrix}\Lambda_1^{(s)} & 0\\ 0 & 0\end{pmatrix}V_1,$$

其中 $\Lambda^{(r)}=\mathrm{diag}(\lambda_1,\cdots,\lambda_r)$，$\lambda_1\geqslant\cdots\geqslant\lambda_r>0$，$\Lambda_1^{(s)}=\mathrm{diag}(\mu_1,\cdots,\mu_s)$，$\mu_1\geqslant\cdots\geqslant\mu_s>0$，则有 $s=r$，$\Lambda_1=\Lambda$，且

$$V^{-1}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}U^{-1}=V_1^{-1}\begin{pmatrix}\Lambda_1^{-1} & 0\\ 0 & 0\end{pmatrix}U_1^{-1}$$

这证明了广义逆矩阵之定义与酉方阵 U,V 之选取无关.

证 由 $r=\text{rank}(A), s=\text{rank}(A)$，所以 $r=s$. 又由

$$\bar{U}_1'U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}\Lambda_1 & 0\\ 0 & 0\end{pmatrix}V_1\bar{V}',$$

将 $\bar{U}_1'U, V_1\bar{V}'$ 都按前 r 行，前 r 分成四块：

$$\bar{U}_1'U=\begin{pmatrix}B & C\\ D & E\end{pmatrix},\quad V_1\bar{V}'=\begin{pmatrix}P & Q\\ R & S\end{pmatrix}.$$

于是有

$$\begin{pmatrix}B\Lambda & 0\\ D\Lambda & 0\end{pmatrix}=\begin{pmatrix}\Lambda_1P & \Lambda_1Q\\ 0 & 0\end{pmatrix}.$$

注意到 $\det\Lambda\neq0$, $\det\Lambda_1\neq0$，所以有 $D=0, Q=0$. 但是 $\bar{U}_1'U, \bar{V}_1\bar{V}$ 都是酉方阵，所以有 $C=0, R=0$. 即有

$$\bar{U}_1'U=\begin{pmatrix}B & 0\\ 0 & E\end{pmatrix},\ V_1\bar{V}'=\begin{pmatrix}P & 0\\ 0 & S\end{pmatrix},$$

且有 $B\Lambda=\Lambda_1P$. 今 P, B 都是 r 阶酉方阵，所以由 $P=\Lambda_1^{-1}B\Lambda$ 有 $\Lambda_1^{-1}B\Lambda^2\bar{B}'\Lambda_1^{-1}=I$，即 $B\Lambda^2\bar{V}'=\Lambda_1^2$. 这证明了 $\Lambda^2=\Lambda_1^2$，即 $\Lambda=\Lambda_1$，且 $B\Lambda=\dot{\Lambda}_1B$. 由 $B\Lambda=\Lambda_1P$ 可知 $P=B, B\Lambda=\Lambda B$，而

$$\begin{aligned}\bar{V}'\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\bar{U}'&=\bar{V}_1'\begin{pmatrix}P & 0\\ 0 & S\end{pmatrix}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}\bar{B}' & 0\\ 0 & \bar{E}'\end{pmatrix}\bar{U}_1'\\ &=\bar{V}_1'\begin{pmatrix}P\Lambda^{-1}\ \ B^{-1} & 0\\ 0 & 0\end{pmatrix}\bar{U}_1'\\ &=\bar{V}_1'\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\bar{U}_1'.\end{aligned}$$

这证明了引理. 证完.

定理12.1.1 线性方程组

$$Ax=\beta$$

的最小二乘通解为

$$x=A^+\beta+(I^{(m)}-A^+A)\xi,\quad \forall\xi\in C^m$$

所以模最小解为

$$x_0=A^+\beta.$$

证 令 $A=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V$, 于是 $\bar{A}'Ax=\bar{A}'\beta$ 可改写为

$$\bar{V}'\begin{pmatrix}\Lambda^2 & 0\\ 0 & 0\end{pmatrix}Vx=\bar{V}'\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}\bar{U}'\beta.$$

记 $y=Vx=\begin{pmatrix}u\\ v\end{pmatrix}$, $u\in V_r$, $v\in V_{m-r}$, 则有

$$\begin{pmatrix}\Lambda^2 u\\ 0\end{pmatrix}=\begin{pmatrix}(\Lambda\ 0)\bar{U}'\beta\\ 0\end{pmatrix},$$

此即

$$u=\Lambda^{-2}(\Lambda\ 0)\bar{U}'\beta=(\Lambda^{-1}0)\bar{U}'\beta.$$

于是

$$x=\bar{V}'\begin{pmatrix}u\\ v\end{pmatrix}=\bar{V}'\begin{pmatrix}(\Lambda^{-1}0)\bar{U}'\beta\\ v\end{pmatrix}=\bar{V}'\begin{pmatrix}(\Lambda^{-1}0)\bar{U}'\beta\\ 0\end{pmatrix}+\bar{V}'\begin{pmatrix}0\\ v\end{pmatrix}.$$

所以

$$x=\bar{V}'\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\bar{U}'\beta+\bar{V}'\begin{pmatrix}0\\ v\end{pmatrix}.$$

令

$$I-A^+A=I-\bar{V}'\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix}V=\bar{V}'\begin{pmatrix}0 & 0\\ 0 & I^{(m-r)}\end{pmatrix}V,$$

这证明了 $x=A^+\beta+(I-A^+A)\xi$, $\forall\xi\in V_m$. 即求出了通解.

下面计算解的模. 由于

$$\overline{(I-A^+A)}'\ A^+=\bar{V}'\begin{pmatrix}0 & 0\\ 0 & I\end{pmatrix}V\bar{V}'\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\bar{U}'=0,$$

所以

$$\begin{aligned}|A^+\beta+(I-A^+A)\xi|^2&=\overline{(A^+\beta+(I-A^+A)\xi)}'\,(A^+\beta+(I-A^+A)\xi)\\ &=|A^+\beta|^2+|(I-A^+A)\xi|^2\geqslant|A^+\beta|^2.\end{aligned}$$

且当 $\xi=0$ 时等式成立. 这证明了 $A^{+}\beta$ 为模最小解. 证完.

下面证明当线性方程组 $Ax=\beta$ 相容时，最小二乘通解就是它在通常意义下的通解，所以最小二乘解是在不相容线性方程组的情形的推广.

定理 12.1.2 线性方程组

$$Ax=\beta$$

相容当且仅当

$$AA^{+}\beta=\beta.$$

这时最小二乘解就是通解.

证 今 $Ax=\beta$ 可改写为 $U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}Vx=\beta$，即

$$\begin{pmatrix}\Lambda^{(r)} & 0\\ 0 & 0\end{pmatrix}Vx=\bar{U}'\beta.$$

记 $Vx=y$，记 $y=\begin{pmatrix}u\\ v\end{pmatrix}$，$u\in V_r$，$v\in V_{m-r}$，则上式可写为

$$\begin{pmatrix}\Lambda u\\ 0\end{pmatrix}=\bar{U}'\beta.$$

于是线性方程组 $Ax=\beta$ 有解，当且仅当 $\bar{U}'\beta=\begin{pmatrix}\alpha\\ 0\end{pmatrix}$，其中 $\alpha\in V_r$，且 $\Lambda u=\alpha$，即 $u=\Lambda^{-1}\alpha$. 因此通解为

$$\begin{aligned}x=\bar{V}'y=\bar{V}'\begin{pmatrix}u\\ v\end{pmatrix}&=\bar{V}'\begin{pmatrix}\Lambda^{-1}\alpha\\ v\end{pmatrix}=\bar{V}'\begin{pmatrix}\Lambda^{-1}\alpha\\ 0\end{pmatrix}+\bar{V}'\begin{pmatrix}0\\ v\end{pmatrix}\\ &=\bar{V}'\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}\alpha\\ 0\end{pmatrix}+\bar{V}'\begin{pmatrix}0\\ v\end{pmatrix}\\ &=\bar{V}'\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\bar{U}'\beta+\bar{V}'\begin{pmatrix}0\\ v\end{pmatrix}.\end{aligned}$$

这证明了

$$x=A^{+}\beta+(I-A^{+}A)\xi,\quad \forall\xi\in V_m.$$

由定理 12.1.1 便证明了 $Ax=\beta$ 之通解也就是最小二乘通解.

上面实际上证明了 $Ax=\beta$ 有解当且仅当 $\overline{U}'\beta=\begin{pmatrix}\alpha\\0\end{pmatrix}$, 其中 $\alpha\in V_r$. 今

$$I-AA^{+}=I-U\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}V\overline{V}'\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\overline{U}'=U\begin{pmatrix}0 & 0\\0 & I\end{pmatrix}\overline{U}',$$

所以 $Ax=\beta$ 有解当且仅当 $(I-AA^{+})\beta=0$, 此即 $AA^{+}\beta=\beta$. 证完.

下面给出强广义逆矩阵的两种等价定义

定理 12.1.3 给定 $n\times m$ 复矩阵 A, 则矩阵方程组

$$AXA=A,\ XAX=X,\ \overline{(AX)'}=AX,\ \overline{(XA)'}=XA$$

有唯一解, 其中 X 为由 mn 个独立自变量构成的 $m\times n$ 矩阵. 这个解就是 A 的强广义逆矩阵 A^{+}.

证 今

$$A=U\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}V,$$

符号意义同上, 于是方程组变为

$$U\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}VXU\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}V=U\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}V,\quad XU\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}VX=X,$$

$$\overline{X}'\overline{V}'\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}\overline{U}'=U\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}VX,\ \overline{V}'\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}\overline{U}'\overline{X}'=XU\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}V.$$

令 $Y=VXU$, 则有

$$\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}Y\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}=\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix},\ Y\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}Y=Y,$$

$$\overline{Y}'\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}=\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}Y,\ Y\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}=\begin{pmatrix}\Lambda & 0\\0 & 0\end{pmatrix}\overline{Y}'.$$

将 Y 按前 r 行及前 r 列分成四块, 记作

$$Y=\begin{pmatrix}P & Q\\ R & S\end{pmatrix}.$$

于是上面矩阵方程组可改写为

$$\Lambda P\Lambda=\Lambda,\ P\Lambda P=P,\ P\Lambda Q=Q,\ R\Lambda P=R, R\Lambda Q=S,$$

$$\bar{P}'\Lambda=\Lambda P, \Lambda Q=0,\ \bar{Q}'\Lambda=0, P\Lambda=\Lambda\bar{P}',\ R\Lambda=0, \Lambda\bar{R}'=0.$$

此即 $Q=0, R=0, P=\Lambda^{-1}, S=0$. 所以解唯一存在，它是

$$X=\bar{V}'Y\bar{U}'=\bar{V}'\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\bar{U}'=A^{+}.$$

证完.

上面定理给出可以用矩阵方程组的解来定义强广义逆矩阵，下面用几何语言，给出强广义逆矩阵的几何定义.

引理 12.1.5 给定 m 维酉空间 $\mathfrak{L}$ 和 n 维酉空间 $\mathfrak{R}$. 设 $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{R}$ 内的线性映射，则唯一存在 $\mathfrak{R}$ 到 $\mathfrak{L}$ 内的线性映射 $\mathscr{A}^*$，它有

$$(\mathscr{A}(u), v)_{\mathfrak{R}}=(u, \mathscr{A}^*(v))_{\mathfrak{L}}, \forall\, u\in\mathfrak{L}, v\in\mathfrak{R}.$$

这时 $\mathscr{A}^*$ 称为 $\mathscr{A}$ 的共轭映射.

证 在 $\mathfrak{L}$ 及 $\mathfrak{R}$ 中分别取标准正交基 $e_1, e_2, \cdots, e_m$ 及 $f_1, f_2, \cdots, f_n$. 于是

$$\mathscr{A}(e_i)=\sum_{j=1}^{n}a_{ji}f_j, i=1,2,\cdots, m$$

它决定一个 $n\times m$ 矩阵 $A=(a_{ij})$，称为 $\mathscr{A}$ 的矩阵表示. 若 $\mathscr{A}^*$ 存在，则有

$$\mathscr{A}(^*f_j)=\sum_{i=1}^{m}b_{ij}e_i, j=1,2,\cdots,n$$

由定义条件，$(\mathscr{A}(e_i), f_j)=(e_i, \mathscr{A}^*(f_j))$，即

$$\left(\sum_{k=1}^{n}a_{ki}f_k,\ f_j\right)=\left(e_i,\ \sum_{l=1}^{m}b_{lj}e_l\right).$$

但是酉空间的内积为 Hermite 双线性函数，所以有

$$a_{ij}=\bar{b}_{ji},\ i=1,2,\cdots,n,\ j=1,2,\cdots,m.$$

因此当$\mathscr{A}$的矩阵表示为A时，$\mathscr{A}^*$的矩阵表示必为$\bar{A}'$. 这也证明了对任一线性映射$\mathscr{A}:\mathfrak{L}\to\mathscr{R}$，则唯一存在它的共轭映射$\mathscr{A}^*:\mathfrak{R}\to\mathfrak{L}$. 证完.

引理 12.1.6 酉空间$\mathfrak{L}$上线性变换$\mathscr{R}$为**正交投影**当且仅当$\mathscr{R}$在象空间$\mathfrak{L}_1=\mathscr{R}(\mathfrak{L})$上为恒等变换，在其正交补$\mathfrak{L}_1^{\perp}$上为零变换；当且仅当$\mathscr{R}^2=\mathscr{R}$, $\mathscr{R}^*=\mathscr{R}$.

证 今$\mathscr{R}$为正交投影，即存在子空间$\mathfrak{L}_1$.对直接和$\mathfrak{L}=\mathfrak{L}_1\oplus\mathfrak{L}_1^{\perp}$，任取$\alpha\in\mathfrak{L}$,则有唯一分解$\alpha=\beta+\gamma$，其中$\beta\in\mathfrak{L}_1$，$\gamma\in\mathfrak{L}_1^{\perp}$. 而$\mathscr{R}:\alpha\to\beta$. 这证明了$\mathscr{R}(\beta)=\beta$，$\mathscr{R}(\gamma)=0$. 即$\mathscr{R}$在$\mathfrak{L}_1$上为恒等变换，在$\mathfrak{L}_1^{\perp}$上为零变换. 反之，则

$$\mathscr{R}(\alpha)=\mathscr{R}(\beta+\gamma)=\mathscr{R}(\beta)+\mathscr{R}(\gamma)=\beta,\ \forall\alpha\in\mathfrak{L}.$$

所以$\mathscr{R}$为正交投影.

再若$\mathscr{R}$为正交投影，自然

$$\mathscr{R}^2(\alpha)=\mathscr{R}(\mathscr{R}(\alpha)=\mathscr{R}(\beta)=\beta=\mathscr{R}(\alpha),\ \forall\alpha\in\mathfrak{L},$$

其中$\alpha=\beta+\gamma$，$\beta\in\mathfrak{L}_1$，$\gamma\in\mathfrak{L}_1^{\perp}$. 于是$\mathscr{R}^2=\mathscr{R}$. 又任取$\alpha=\beta+\gamma$，$\xi=\eta+\zeta$，其中$\alpha,\zeta\in\mathfrak{L},\beta$，$\eta\in\mathfrak{L}_1,\gamma$，$\zeta\in\mathfrak{L}_1^{\perp}$. 于是由共轭变换之定义，有$(\mathscr{R}(\alpha),\xi)=(\alpha,\ \mathscr{R}^*(\xi))$. 因此记$\mathscr{R}^*(\xi)=\lambda+\mu$，其中$\lambda\in\mathfrak{L}_1$, $\mu\in\mathfrak{L}_1^{\perp}$，则有$(\beta,\xi)=(\alpha,\lambda+\mu)$，于是$(\beta,\eta)=(\beta,\lambda)+(\gamma,\mu)$. 由$\alpha$任取，我们取$\alpha\in\mathfrak{L}_1$，则$\alpha=\beta,\gamma=0$，所以$(\beta,\eta-\lambda)=0$. 这证明了$\eta=\lambda$. 因此对一般的$\alpha\in\mathfrak{L}_1$，则可任取$\gamma\in\mathfrak{L}_1^{\perp}$，使得$(\gamma,\mu)=0$，其中$\gamma,\mu\in\mathfrak{L}_1^{\perp}$. 这证明了$\mu=0$. 所以$\mathscr{R}^*(\xi)=\mathscr{R}^*(\eta+\zeta)=\lambda$ 这证明了$\mathscr{R}^*(\alpha)=\beta,\forall\alpha\in\mathfrak{L}$，即$\mathscr{R}^*=\mathscr{R}$. 反之，若$\mathfrak{L}$上线性变换$\mathscr{R}$有$\mathscr{R}^2=\mathscr{R},\mathscr{R}^*=\mathscr{R}$. 由共轭线性变换之定义可知，在$\mathfrak{L}$中存在标准正交基$\alpha_1,\alpha_2,\cdots,\alpha_n$，使得$\mathscr{R}$之方阵表示$P$有$P^2=P$，$\bar{P}'=P$. 所以$P$酉相似于对角形$\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$，且有$\lambda_j^2=\lambda_j$，即$\lambda_j=0$或$\lambda_j=1$. 所以证明了在$\mathfrak{L}$中存在一组标准正交基$\beta_1,\beta_2,\cdots,\beta_n$，使得$\mathscr{R}$之方阵

表示为$\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}$, 即$\mathscr{R}(e_i)=e_i$, $1\leqslant i\leqslant r$; $\mathscr{R}(e_j)=0, r+1\leqslant j\leqslant n$. 记以$e_1,\cdots,e_r$为基之子空间为$\mathfrak{L}_1$, 则$\mathscr{R}$在$\mathfrak{L}_1$上为恒等映射, 在$\mathfrak{L}_1^{\perp}$上为零映射. 所以$\mathscr{R}$为正交投影. 证完.

定义 设$\mathscr{A}$为m维酉空间$\mathfrak{L}$到n维酉空间$\mathfrak{R}$内的线性映射, 则$\mathscr{R}_{\mathscr{A}}$为$\mathfrak{R}$中关于$\mathscr{A}(\mathfrak{L})$之正交投影.

于是有

引理 12.1.7 在m维酉空间$\mathfrak{L}$中任取标准正交基$e_1, e_2,\cdots, e_m$, 在n维酉空间$\mathfrak{R}$中任取标准正交基$f_1,f_2,\cdots, f_n$. 设$\mathfrak{L}$到$\mathfrak{R}$内的线性映射$\mathscr{A}$有矩阵表示$A=(a_{ij})=U\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}V$, 其中$U$为$n$阶酉方阵, V为m阶酉方阵, $\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_r)$; $\lambda_1\geqslant\cdots\geqslant\lambda_r>0$. 于是线性映射$\mathscr{A}^*:\mathfrak{R}\to\mathfrak{L}$; $\mathscr{R}_{\mathscr{A}}:\mathfrak{R}\to\mathfrak{R}$; $\mathfrak{R}_{\mathscr{A}^*}:\mathfrak{L}\to\mathfrak{L}$分别有矩阵表示

$$\bar{A}'=\bar{V}'\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}\bar{U}',\ P_{\mathscr{A}}=U\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}\bar{U}', P_{\mathscr{A}^*}=\bar{V}'\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}V.$$

证 今存在n阶酉方阵U及m阶酉方阵V, 使得

$$A=U\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}V, \Lambda=\mathrm{diag}(\lambda_1\cdots,\lambda_r), \lambda_1\geqslant\cdots\geqslant\lambda_r>0.$$

记$U=(u_{ij}), V=(v_{ij})$, 于是作标准正交基到标准正交基的基变换

$$e'_i=\sum_{j=1}^{m}\bar{v}_{ij}e_j, 1\leqslant i\leqslant m;\ f'_i=\sum_{j=1}^{n}u_{ji}f_j, 1\leqslant i\leqslant n,$$

则有

$$f_i=\sum_{j=1}^{n}\bar{u}_{ij}f'_j, 1\leqslant i\leqslant n.$$

且

$$\mathscr{A}(e'_i)=\sum\bar{v}_{ij}\mathscr{A}(e_j)=\sum\bar{v}_{ij}a_{kj}f_k=\sum\bar{v}_{ij}a_{kj}\bar{u}_{kl}f'_l,\ 1\leqslant i\leqslant m.$$

这证明了线性映射$\mathscr{A}:\mathfrak{L}\to\mathfrak{R}$在$\mathfrak{L}$的标准正交基$e'_1,\cdots, e'_m$及$\mathfrak{R}$的

标准正交基 $f'_1,\cdots,f'_n$ 下的矩阵表示为 $C=(c_{ij})$，即

$$\mathscr{A}(e'_i)=\sum_{l=1}^{n}c_{li}f'_l,\ 1\leqslant i\leqslant m$$

其中

$$C=\bar{U}'A\bar{V}'=\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}.$$

此即

$$\begin{cases}\mathscr{A}(e'_i)=\lambda_i f'_i,\ 1\leqslant i\leqslant r,\\ \mathscr{A}(e'_j)=0,\ r+1\leqslant j\leqslant m.\end{cases}$$

于是$\mathscr{A}(\mathfrak{L})$是以 $f'_1,\cdots,f'_r$ 为标准正交基的子空间. 由引理 12.1.6，在 $\mathfrak{R}$ 的标准正交基$f'_1,f'_2,\cdots,f'_n$下，$\mathscr{R}_{\mathscr{A}}$有方阵表示$\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix}$. 于是在 $\mathfrak{R}$ 的标准正交基 $f_1,f_2,\cdots,f_n$ 下，$\mathscr{R}_{\mathscr{A}}$有方阵表示

$$P_{\mathscr{A}}=U\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix}\bar{U}'.$$

同理，$\mathscr{A}^*$有矩阵表示 $\bar{A}'=\bar{V}'\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}\bar{U}'$，于是$\mathscr{R}_{\mathscr{A}^*}$有方阵表示

$$P_{\mathscr{A}^*}=\bar{V}'\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix}V.$$

证完.

定义 给定的m维酉空间$\mathfrak{L}$到 n 维酉空间 $\mathfrak{R}$ 内的线性映射 $\mathscr{A}$ 及 $\mathfrak{R}$ 到 $\mathfrak{L}$ 内的线性映射$\mathscr{L}$. 如果有

(1) $\operatorname{rank}(\mathscr{A})=\operatorname{rank}(\mathscr{L})$;

(2) $\mathscr{A}\mathscr{L}=\mathscr{R}_{\mathscr{A}},\mathscr{L}\mathscr{A}=\mathscr{R}_{\mathscr{A}^*}$,

则 $\mathscr{L}$ 称为$\mathscr{A}$的**强广义逆映射**，改记 $\mathscr{L}$ 为$\mathscr{A}^+$.

定理 12.1.4 符号同引理 12.1.7，记 $\mathscr{L}$ 有矩阵表示 B，则$\mathscr{L}$为$\mathscr{A}$的强广义逆映射当且仅当 $B=A^+$.

证 设$\mathscr{L}$为$\mathscr{A}$的强广义逆映射，即$\mathrm{rank}(\mathscr{L})=\mathrm{rank}(\mathscr{A})$，且$\mathscr{A}\mathscr{L}=\mathscr{R}_{\mathscr{A}}$，$\mathscr{L}\mathscr{A}=\mathscr{R}_{\mathscr{A}^*}$. 由引理12.1.7，所以矩阵关系

$$\mathrm{rank}(A)=\mathrm{rank}(B),AB=P_{\mathscr{A}},BA=P_{\mathscr{A}^*}.$$

成立当且仅当

$$\mathrm{rank}(B)=r,\ U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}VB=U\begin{pmatrix}I & 0\\ 0 & 0\end{pmatrix}\overline{U}',$$

$$BU\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V=\overline{V}'\begin{pmatrix}I & 0\\ 0 & 0\end{pmatrix}V.$$

记

$$B_1=\overline{V}'BU=\begin{pmatrix}D^{(r)} & E\\ F & G\end{pmatrix},$$

则当且仅当

$$\mathrm{rank}(B_1)=r,\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}D & E\\ F & G\end{pmatrix}=\begin{pmatrix}I & 0\\ 0 & 0\end{pmatrix},$$

$$\begin{pmatrix}D & E\\ F & G\end{pmatrix}\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}I & 0\\ 0 & 0\end{pmatrix}.$$

此即$D=\Lambda^{-1},E=0,F=0$. 于是$B_1=\overline{V}BU=\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & G\end{pmatrix}$.

由于$\mathrm{rank}(B_1)=r$，所以$G=0$. 这证明了当且仅当

$$B=\overline{V}'\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\overline{U}'=A^+.$$

证完.

将上面的理论全部限制在实的情形，即改酉空间为Euclid空间，酉方阵为实正交方阵，则全部结论都成立.

习题 12.1

1. 设A为$n\times m$复矩阵. 试证：$(A^+)^+=A$，$(\overline{A}')^+=(\overline{A}^+)'$，$\overline{A}'(\overline{A}'A\overline{A}')^+\overline{A}'=A^+$. 又当$A$为$n$阶非异方阵，则$A^+=A^{-1}$.

2. 设A为$n\times m$矩阵，P为n阶酉方阵，Q为m阶酉方阵，则$(PAQ)^+=$

$Q^{-1}A^{+}P^{-1}$.

3. 设 A 为 $n\times m$ 矩阵，B 为 $m\times q$ 矩阵，且 $AB=0$. 试证：$A(\bar{B}')^{+}=0$，$(\bar{A}')^{+}B=0$.

4. 试举反例来说明：当 A 为 $n\times m$ 矩阵，B 为 $m\times q$ 矩阵，则一般来说，$(AB)^{+}=B^{+}A^{+}$不成立.

5. 试给出 $n\times 1$ 矩阵及 $1\times m$ 矩阵的强广义逆矩阵的明显表达式.

6. 设 A 为 $n\times m$ 矩阵. 试证：存在 $n\times r$ 矩阵 B 及 $r\times m$ 矩阵 C，使得 $\mathrm{rank}(B)=\mathrm{rank}(C)=r$，其中 $r=\mathrm{rank}(A)$. 又 $A=BC$，且

$$A^{+}=\bar{C}'(C\bar{C}')^{-1}(\bar{B}'B)^{-1}\bar{B}'=C^{+}B^{+}.$$

§12.2 广义逆矩阵

定义 设 X 为由 mn 个独立未知数构成的 $m\times n$ 矩阵，A 为 $n\times m$ 矩阵，则矩阵方程

$$AXA=A$$

的解称为 A 的**广义逆矩阵**，记作 A^{-}.

注意：矩阵方程 $AXA=A$ 之解不唯一，所以 A^{-}是泛指其中一个解.

定理 12.2.1 设 A 为 $n\times m$ 复矩阵，且

$$A=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V,\ \Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_r),\ \lambda_1\geqslant\cdots\geqslant\lambda_r>0,$$

其中 U 及 V 分别为 n 阶及 m 阶酉方阵. 则 A 的广义逆矩阵

$$A^{-}=\bar{V}'\begin{pmatrix}\Lambda^{-1} & X_{12}\\ X_{21} & X_{22}\end{pmatrix}\bar{U}'=A^{+}+\bar{V}'\begin{pmatrix}0 & X_{12}\\ X_{21} & X_{22}\end{pmatrix}\bar{U}'.$$

其中 X_{12},X_{21},X_{22}为任意矩阵. 所以广义逆矩阵不唯一，实际上它们全体构成一个集合，而 A^{-} 只表示这个集合中的任意一个元素.

证 令 $A=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V$. 记 $X=\bar{V}'\begin{pmatrix}X_{11} & X_{12}\\ X_{21} & X_{22}\end{pmatrix}\bar{U}'$，其中 X_{11} 为 r 阶方阵. 于是 $AXA=A$ 变成

$$\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}$$

此即 $X_{11}=\Lambda^{-1}$. 证完.

定理 12.2.2 设 A 为 $n\times m$ 复矩阵, β 为 $n\times 1$ 复矩阵, x 为由 m 个独立未知数构成的 $m\times 1$ 复矩阵. 则线性方程组

$$Ax=\beta$$

相容的必要且充分条件为

$$AA^{-}\beta=\beta.$$

这时通解为

$$x=A^{-}\beta+(I-A^{-}A)\xi, \forall\xi\in V_m.$$

证 用定理 12.2.1 的符号, 于是 $A=U\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}V$. 因此 $Ax=\beta$ 可改变为 $U\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}Vx=\beta$, 它等价于线性方程组 $\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}Vx=\bar{U}'\beta$. 记 $Vx=y=\begin{pmatrix} u \\ v \end{pmatrix}$, u 为 $r\times 1$ 矩阵. 记 $\bar{U}'\beta=\begin{pmatrix} \sigma \\ \tau \end{pmatrix}$, $\sigma\in V_r$, $\tau\in V_{m-r}$. 于是有

$$\Lambda u=\sigma, \quad \tau=0$$

所以通解为

$$x=\bar{V}'y=\bar{V}'\begin{pmatrix} u \\ v \end{pmatrix}=\bar{V}'\begin{pmatrix} \Lambda^{-1}\sigma \\ v \end{pmatrix}.$$

又 $Ax=\beta$ 相容的必要且充分条件为 $\tau=0$, 即 $\beta=U\begin{pmatrix} \sigma \\ 0 \end{pmatrix}$.

另一方面, 由定理 12.2.1 有

$$A^{-}=\bar{V}'\begin{pmatrix} \Lambda^{-1} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}\bar{U}'.$$

于是

$$I-A^-A=\bar{V}\begin{pmatrix}0 & 0\\ -X_{21}\Lambda & I\end{pmatrix}V,\quad I-AA^-=U\begin{pmatrix}0 & -\Lambda X_{12}\\ 0 & I\end{pmatrix}\bar{U}'.$$

于是

$$(I-AA^-)\beta=U\begin{pmatrix}0 & -\Lambda X_{12}\\ 0 & I\end{pmatrix}\bar{U}'U\begin{pmatrix}\sigma\\ \tau\end{pmatrix}=U\begin{pmatrix}-\Lambda X_{12}\tau\\ \tau\end{pmatrix}.$$

所以 $\tau=0$ 当且仅当 $(I-AA^-)\beta=0$，即 $AA^-\beta=\beta$. 又这时通解为

$$x=\bar{V}'\begin{pmatrix}\Lambda^{-1}\sigma\\ v\end{pmatrix}=\bar{V}'\begin{pmatrix}\Lambda^{-1}\sigma\\ 0\end{pmatrix}+\bar{V}'\begin{pmatrix}0\\ v\end{pmatrix}$$

$$=\bar{V}'\begin{pmatrix}\Lambda^{-1} & X_{12}\\ X_{21} & X_{22}\end{pmatrix}\begin{pmatrix}\sigma\\ \tau\end{pmatrix}-\bar{V}'\begin{pmatrix}0\\ X_{21}\sigma\end{pmatrix}+\bar{V}'\begin{pmatrix}0\\ v\end{pmatrix}$$

$$=A^-\beta+(I-A^-A)\xi,\ \forall\xi\in V_m.$$

证完.

定理 12.2.3 线性函数组 $f(x)=Ax-\beta$ 的最小二乘通解为

$$x=(\bar{A}'A)^-\bar{A}'\beta+[I-(\bar{A}'A)^-(\bar{A}'A)]\xi,\quad \forall\xi\in V_m$$

证 由引理 12.1.1, $f(x)=Ax-\beta$ 的最小二乘通解为相容线性方程组 $\bar{A}'Ax=\bar{A}'\beta$ 的通解. 由定理 12.2.2 便证明了定理. 证完.

广义逆矩阵可用来求某些矩阵方程之通解.

定理 12.2.4 给定 $n\times m, m\times p, p\times q, n\times q$ 矩阵 A, X, B, C，其中 A, B, C 为常数矩阵，X 为由 mp 个独立未知数构成的矩阵. 则矩阵方程

$$AXB=C$$

有解当且仅当

$$AA^-CB^-B=C,$$

这时通解为

$$X=A^-CB^-+Z-A^-AZBB^-,$$

其中 Z 取任一 $m\times p$ 复矩阵.

证 记 $A=(a_{ij})$，$X=(x_{jk})$，$B=(b_{kl})$，$c=(c_{il})$. 于是矩阵

方程可写为

$$\sum a_{ij}x_{jk}b_{kl}=c_{il},\ 1\leqslant i\leqslant n,\ 1\leqslant l\leqslant q.$$

将 X,C 中元素按行依次排为 $1\times mp$, $1\times nq$ 矩阵

$$x=(x_{11},\cdots,x_{1p},\cdots,x_{m1},\cdots,x_{mp}).$$

$$c=(c_{11},\cdots,c_{1q},\cdots,c_{n1},\cdots,c_{nq}).$$

将上面矩阵方程用线性方程组的通常方式来表达，即有

$$Dx'=c',$$

于是 D 为 $nq\times mp$ 矩阵，它可表为

$$D=A\otimes B',$$

其中$\otimes$称为矩阵的**Kronecker 乘积**，定义为：设 $A=(a_{ij})$ 为 $n\times m$ 矩阵，B 为 $p\times q$ 矩阵，则

$$A\otimes B=\begin{pmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{pmatrix}.$$

由定义出发，不难证明

(1) $(A\otimes B)'=A'\otimes B'$,

(2) $\overline{(A\otimes B)}=\bar{A}\otimes\bar{B}$,

(3) $(A_1+A_2)\otimes B=A_1\otimes B+A_2\otimes B$,

(4) $A\otimes(B_1+B_2)=A\otimes B_1+A\otimes B_2$,

(5) $(\lambda A)\otimes(\mu B)=\lambda\mu(A\otimes B),\ \forall\lambda,\mu\in C$

(6) $(A_1\otimes B_1)(A_2\otimes B_2)=A_1A_2\otimes B_1B_2$,

(7) $(A\otimes B)^+=A^+\otimes B^+$,

(8) A 的任一广义逆矩阵 A^- 和 B 的任一广义逆矩阵 B^- 的 Kronecker 乘积 $A^-\otimes B^-$ 必为 $(A\otimes B)$ 的广义逆矩阵. 反之不一定.

下面用这些性质来解矩阵方程 $AXB=C$. 事实上，今 $(A\otimes B')x'=c'$. 由定理 12.2.2，它有解的必要且充分条件为对 $A\otimes B'$ 的任意一个广义逆矩阵 $(A\otimes B')^-$，则有

$$(A\otimes B')(A\otimes B')^{-}c'=c'.$$

这时通解为

$$x'=(A\otimes B')^{-}c'+\xi-(A\otimes B')^{-}(A\otimes B')\xi.$$

由性质(8)，$A^{-}\otimes(B')^{-}=A^{-}\otimes(B^{-})'$为$A\otimes B'$的广义逆矩阵．于是取$(A\otimes B')^{-}$为$A^{-}\otimes(B')^{-}$，则有

$$(A\otimes B')(A^{-}\otimes(B^{-})')c'=(AA^{-}\otimes(B^{-}B)')c'=c',$$

且通解为

$$x'=(A^{-}\otimes(B^{-})')c'+\xi-(A^{-}\otimes(B^{-})')(A\otimes B')\xi.$$

写回矩阵形式，即有解之充分且必要条件为

$$AA^{-}CB^{-}B=C,$$

又通解为

$$X=A^{-}CB^{-}+Z-A^{-}AZBB^{-},$$

其中 Z 为 ξ'排成的 $m\times p$ 矩阵．证完．

第十三章　方阵在相似下的标准形

§13.1　λ矩阵在相抵下的标准形

定义　元素由未知数λ的多项式构成的矩阵称为λ矩阵.

回忆第二、四章可知，只要不考虑n阶非异方阵A的逆方阵$A^{-1}=(\det A)^{-1}A^*$,矩阵的所有定义和性质对λ矩阵都成立. 这是因为一般来说$\det A$为λ的多项式,除非$\det A$除得尽伴随方阵A^*中所有元素. 特别当$\det A$为非零常数时，λ矩阵A^{-1}就有定义.

定义　设A为n阶λ方阵. 如果存在n阶λ方阵B使得

$$AB=BA=I^{(n)},$$

则A称为**可逆的**,这时B称为A的**逆方阵**,记作

$$B=A^{-1}.$$

定理 13.1.1　n阶λ方阵A可逆当且仅当$\det A$为非零常数，这时

$$A^{-1}=(\det A)^{-1}A^*.$$

证　设A可逆,即存在λ方阵B，使得$AB=BA=I^{(n)}$. 于是$1=(\det A)(\det B)$. 由于$\det A,\det B$都是λ的多项式，这证明了$0=\deg 1=\deg(\det A)+\deg(\det B)$.所以$\deg(\det A)=0$,即$\det A$为非零常数.反之,若$\det A$为非零常数,由于Laplace展开定理可知$AA^*=A^*A=(\det A)I$.因此$A^{-1}=(\det A)^{-1}A^*$为$A$的逆方阵.证完.

又初等变换也可以推广到λ矩阵的情形

定义　设A为$n\times m\lambda$矩阵.将A的第i行(列)和第j行(列)互换，称为对A的行(列)作了**第一类行(列)的初等变换**. 将A的第i行(列)乘以非零常数，称为对A的行(列)作了**第二类行(列)的初等变换**. 将A的第i行(列)乘以多项式$f(\lambda)$，再加到第j行

(列),称为对 A 的行(列)作了**第三类行(列)的初等变换.**

定义 n 阶 λ 方阵

$$P_{ij}=\begin{pmatrix} 1 &&&&&&&&& \\ & \ddots &&&&&&&& \\ && 1 &&&&&&& \\ &&& 0 & \cdots & 1 &&&& \\ &&&& 1 &&&&& \\ &&& \vdots & \ddots & \vdots &&&& \\ &&&&& 1 &&&& \\ &&& 1 & \cdots & 0 &&&& \\ &&&&&& 1 &&& \\ &&&&&&& \ddots && \\ &&&&&&&& 1 & \end{pmatrix}\begin{matrix} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{matrix},\quad 1\leqslant i<j\leqslant n$$

$i\qquad j$

称为**第一类初等 λ 方阵.** n 阶 λ 方阵

$$P_i(a)=\mathrm{diag}(1,\cdots,1,\overset{(i)}{a},1,\cdots,1),1\leqslant i\leqslant n$$

称为**第二类初等 λ 方阵**,其中 a 为非零常数. n 阶方阵

$$Q_{ji}(f(\lambda))=\begin{pmatrix} 1 &&&&&&&&& \\ & \ddots &&&&&&&& \\ && 1 &&&&&&& \\ &&& 1 & \cdots & 0 &&&& \\ &&&& 1 &&&&& \\ &&& \vdots & \ddots & \vdots &&&& \\ &&&&& 1 &&&& \\ &&& f(\lambda) & \cdots & 1 &&&& \\ &&&&&& 1 &&& \\ &&&&&&& \ddots && \\ &&&&&&&& 1 & \end{pmatrix}\begin{matrix} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{matrix},1\leqslant i\neq j\leqslant n,$$

$i\qquad j$

称为**第三类初等 λ 方阵**，其中 $f(\lambda)$ 为 λ 的多项式.

引理 13.1.1 设 A 为 $n\times m\lambda$ 矩阵. 对 A 的行(列)作第一、二、三类行(列)的初等变换，等于在 A 的左边(右边)分别乘以第一、二、三类初等 λ 方阵.

引理 13.1.2 初等 λ 方阵为可逆 λ 方阵，且

$$P_{ij}^{-1}=P_{ij}, P_i(a)^{-1}=P_i(a^{-1}), Q_{ij}(f(\lambda))^{-1}=Q_{ij}(-f(\lambda)).$$

所以初等 λ 方阵的逆方阵是同类型的初等 λ 方阵.

这两个引理的证明是显然的.

对 λ 矩阵，下面定义是非常重要的.

定义 设 A 为 $n\times m\lambda$ 矩阵，$r=\mathrm{rank}(A)$. A 中所有 t 阶子式的首项系数为 1 的最大公因式 $D_t(x)$ 称为 A 的 t 阶**行列式因式**，其中 $t=1,2,\cdots,r$.

引理 13.1.3 设 A 为 $n\times m\lambda$ 矩阵，秩为 r. 记 $D_1(\lambda), D_2(\lambda), \cdots, D_r(t)$ 为 A 的行列式因式. 则有

$$D_1(\lambda)\mid D_2(\lambda), D_2(\lambda)\mid D_3(\lambda), \cdots, D_{r-1}(\lambda)\mid D_r(\lambda).$$

证 由 Laplace 展开定理及最大公因式定义可知引理成立. 证完.

定义 设 A 为 $n\times m\lambda$ 矩阵，$\mathrm{rank}(A)=r$. $D_1(\lambda),\cdots, D_r(\lambda)$ 为 A 的行列式因式. 则首项系数为 1 的多项式

$$d_1(\lambda)=D_1(\lambda), d_2(\lambda)=\frac{D_2(\lambda)}{D_1(\lambda)}, \cdots, d_r(\lambda)=\frac{D_r(\lambda)}{D_{r-1}(\lambda)}$$

称为 A 的**不变因式**，其中 $d_i(\lambda)$ 称为 A 的第 i 个不变因式.

注意不变因式由行列式因式唯一确定，反之亦然. 事实上，

$$D_1(\lambda)=d_1(\lambda), D_2(\lambda)=d_1(\lambda)d_2(\lambda), \cdots,$$
$$D_r(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda).$$

引理 13.1.4 $n\times m\lambda$ 矩阵 A 的行列式因式在初等变换下不变，所以不变因式也在初等变换下不变.

证 显然作了第一及第二类初等变换，行列式因式不变. 下面只考虑第三类初等变换，且仅限于讨论对行作第三类初等变换的情形，因为对列的情形证明是一样的. 记

$$A=\begin{pmatrix} a_{11}(\lambda) & \cdots & a_{1m}(\lambda) \\ \vdots & & \vdots \\ a_{n1}(\lambda) & \cdots & a_{nm}(\lambda) \end{pmatrix}.$$

其中 $a_{ij}(\lambda)$ 为 λ 的多项式. 例如

$$B=Q_{21}(f(\lambda))A=\begin{pmatrix} 1 & f(\lambda) & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}\begin{pmatrix} a_{11}(\lambda) & \cdots & a_{1m}(\lambda) \\ \vdots & & \vdots \\ a_{n1}(\lambda) & \cdots & a_{nm}(\lambda) \end{pmatrix}$$

$$=\begin{pmatrix} a_{11}(\lambda)+f(\lambda)a_{21}(\lambda) & \cdots & a_{1m}(\lambda)+f(\lambda)a_{2m}(\lambda) \\ a_{21}(\lambda) & \cdots & a_{2m}(\lambda) \\ \vdots & & \vdots \\ a_{n1}(\lambda) & \cdots & a_{nm}(\lambda) \end{pmatrix},$$

于是

$$B\binom{1\, i_2 \cdots i_t}{j_1 j_2 \cdots j_t}=A\binom{1\, i_2 \cdots i_t}{j_1 j_2 \cdots j_t}+f(t)A\binom{2\, i_2 \cdots i_t}{j_1 j_2 \cdots j_t},\ 1<i_2<\cdots<i_t\leqslant n,$$

$$B\binom{i_1 i_2 \cdots i_t}{j_1 j_2 \cdots j_t}=A\binom{i_1 i_2 \cdots i_t}{j_1 j_2 \cdots j_t},\ 1<i_1<i_2<\cdots<i_t\leqslant n,$$

又 $1\leqslant j_1<\cdots<j_t\leqslant m$. 记 $\tilde{D}_t(\lambda)$ 为 B 的 t 阶行列式因式，于是 $\tilde{D}_t(\lambda)\mid A\binom{i_1 i_2 \cdots i_t}{j_1 j_2 \cdots j_t}$，$1<i_1<\cdots<i_t\leqslant n$，$1\leqslant j_1<\cdots<j_t\leqslant m$，因此 $\tilde{D}_t(\lambda)\mid A\binom{1\, i_2 \cdots i_t}{j_1 j_2 \cdots j_t}$，$1<i_2<\cdots<i_t\leqslant n$，$1\leqslant j_1<\cdots<j_t\leqslant m$. 所以 $\tilde{D}_t(\lambda)\mid D_t(\lambda)$. 由于 $A=Q_{21}(f(\lambda))^{-1}B=Q_{21}(-f(\lambda))B$，同理可证 $D_t(\lambda)\mid\tilde{D}_t(\lambda)$. 这证明了 $\tilde{D}_t(\lambda)=D_t(\lambda)$，$t=1,2,\cdots,\mathrm{rank}(A)$. 证完.

定理 13.1.2 设 $r=\mathrm{rank}(A)$ 为 $n\times m\lambda$ 矩阵A的秩. 记$d_1(\lambda)$，$d_2(\lambda),\cdots,d_r(\lambda)$为 A 的 r 个不变因式. 则对 A 作一系列初等变换可将 A 变为标准形

$$\begin{pmatrix} \begin{pmatrix} d_1(\lambda) & & \\ & \ddots & \\ & & d_r(\lambda) \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix}.$$

且不变因式有

$$d_1(\lambda) \mid d_2(\lambda), \cdots, d_{r-1}(\lambda) \mid d_r(\lambda).$$

证 当 $A=0$，不必讨论．设 $A\neq 0$．于是存在一个多项式 $a_{ij}(\lambda)\neq 0$，其中

$$A=\begin{pmatrix} a_{11}(\lambda) & \cdots & a_{1m}(\lambda) \\ \vdots & & \vdots \\ a_{n1}(\lambda) & \cdots & a_{nm}(\lambda) \end{pmatrix}.$$

考虑 $P_{1i}AP_{1j}=\begin{pmatrix} a_{ij}(\lambda) & * \\ * & * \end{pmatrix}$．所以无妨设 $a_{11}(\lambda)\neq 0$．这时考虑第一行 $a_{11}(\lambda), a_{12}(\lambda), \cdots, a_{1m}(\lambda)$．设 $a_{1k}(\lambda)\neq 0$，其中 $k>1$．作 $a_{1k}(\lambda)=q_k(\lambda)a_{11}(\lambda)+r_{1k}(\lambda)$，则 $r_{1k}(\lambda)=0$ 或者 $r_{1k}(\lambda)\neq 0$，但是 $\deg r_{1k}(\lambda)<\deg a_{11}(\lambda)$．考虑 $AQ_{1k}(-q_k(\lambda))$，则有

$$AQ_{1k}(-q_k(\lambda))=$$

$$\begin{pmatrix} a_{11}(\lambda) & \cdots & a_{1,k-1}(\lambda) & r_{1k}(\lambda) & a_{1,k+1}(\lambda) & \cdots & a_{1m}(\lambda) \\ a_{21}(\lambda) & \cdots & a_{2,k-1}(\lambda) & \tilde{a}_{2k}(\lambda) & a_{2,k+1}(\lambda) & \cdots & a_{2m}(\lambda) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1}(\lambda) & \cdots & a_{n,k-1}(\lambda) & \tilde{a}_{nk}(\lambda) & a_{n,k+1}(\lambda) & \cdots & a_{nm}(\lambda) \end{pmatrix}.$$

设 $r_{1k}(\lambda)=0$，则 $AQ_{1k}(-q_k(\lambda))$ 的第一行，第 k 列元素为零．设 $r_{1k}(\lambda)\neq 0$，再作 $AQ_{1k}(-q_k(\lambda))P_{1k}=\begin{pmatrix} r_{1k}(\lambda) & * \\ * & * \end{pmatrix}$，由于 $0\leqslant \deg r_{1k}(\lambda)<\deg a_{11}(\lambda)$．所以将 A 变为另一 λ 矩阵，它的第一行第一列元素次数降低．不断地用这个办法作下去．由于原来 $\deg a_{11}(\lambda)$ 有限，用归纳法便证明了 A 可变为 λ 矩阵 $\begin{pmatrix} \tilde{b}_{11}(\lambda) & 0 \\ * & * \end{pmatrix}$．对它的第一

列和第一行同时用同样办法讨论，便证明了 A 可变为 λ 矩阵

$$B=\begin{pmatrix} b_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & b_{22}(\lambda) & \cdots & b_{2m}(\lambda) \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2}(\lambda) & \cdots & b_{nm}(\lambda) \end{pmatrix},$$

其中 $b_{11}(\lambda)\neq 0$.

若存在指标 j,k，其中 $2\leqslant j\leqslant n, 2\leqslant k\leqslant m$，且有 $b_{11}(\lambda)$除不尽 $b_{jk}(\lambda)$. 考虑 λ 矩阵 $Q_{j1}(1)B$，则有

$$Q_{j1}(1)B=\begin{pmatrix} b_{11}(\lambda) & b_{j2}(\lambda) & \cdots & b_{jm}(\lambda) \\ 0 & b_{22}(\lambda) & \cdots & b_{2m}(\lambda) \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2}(\lambda) & \cdots & b_{nm}(\lambda) \end{pmatrix}.$$

和前面一样的办法，又可降低第一行第一列位置元素的次数. 总之，经过有限步，便可将 A 化为 λ 矩阵 B，其中 $b_{11}(\lambda)\neq 0$，且

$$b_{11}(\lambda)\,|\,b_{jk}(\lambda), 2\leqslant j\leqslant n, 2\leqslant k\leqslant m.$$

对 $(n-1)\times(m-1)\lambda$ 矩阵

$$B_1=\begin{pmatrix} b_{22}(\lambda) & \cdots & b_{2m}(\lambda) \\ \vdots & & \vdots \\ b_{n2}(\lambda) & \cdots & b_{nm}(\lambda) \end{pmatrix},$$

作初等变换等于对 $n\times m\lambda$ 矩阵 B 作初等变换. 所以由归纳法便证明了 $n\times m\lambda$ 矩阵可经过一系列初等变换变为标准形

$$\begin{pmatrix} \begin{pmatrix} d_1(\lambda) & & \\ & \ddots & \\ & & d_r(\lambda) \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix},$$

其中 $d_1(\lambda)\cdots d_r(\lambda)\neq 0$，又 $d_1(\lambda),\cdots, d_r(\lambda)$的首项系数为 1，且有 $d_1(\lambda)\,|\,d_2(\lambda),\cdots,d_{r-1}(\lambda)\,|\,d_r(\lambda)$. 由秩的定义可知 $r=\mathrm{rank}(A)$. 由行列式因式之定义可知标准形之行列式因式为

$$D_1(\lambda)=d_1(\lambda), D_2(\lambda)=d_1(\lambda)d_2(\lambda), \cdots, D_r(\lambda)$$
$$=d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda).$$

因此证明了 $d_1(\lambda), \cdots, d_r(\lambda)$ 为标准形的不变因式. 由引理 13.1.4 可知 $d_1(\lambda), \cdots, d_r(\lambda)$ 为 λ 矩阵 A 的不变因式. 证完.

定理 13.1.3 n 阶可逆 λ 矩阵 A 为有限个初等 λ 方阵之乘积.

证 由于作初等变换等于左乘或右乘初等 λ 方阵. 所以定理 13.1.2 证明了存在初等 λ 方阵之乘积 P,Q, 使得

$$PAQ=\mathrm{diag}(d_1(\lambda), \cdots, d_n(\lambda)).$$

由定理 13.1.1, A 为可逆 λ 方阵当且仅当 $\det A$ 为非零常数, 所以秩为 n. 其中 $d_1(\lambda)|d_2(\lambda), \cdots, d_{n-1}(\lambda)|d_n(\lambda)$, 且 $d_1(\lambda), \cdots, d_n(\lambda)$ 为首项系数等于 1 的非零多项式. 但是初等方阵为可逆 λ 方阵, 所以 $\det P, \det Q$ 都是非零常数. 而

$$\det A=(\det P)^{-1}(\det Q)^{-1}d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda).$$

由于 $\det A$ 为非零常数, 所以证明了 $\deg d_j(\lambda)=0, 1\leqslant j\leqslant n$, 由于 $d_j(\lambda)$ 的首项系数为 1, 因此 $d_j(\lambda)=1$, 所以 $PAQ=I^{(n)}$, 即 $A=P^{-1}Q^{-1}$ 为初等 λ 方阵之乘积. 证完.

定义 $n\times m\lambda$ 矩阵 A 和 B 称为**相抵的**, 如果存在 n 阶可逆 λ 矩阵 P 和 m 阶可逆 λ 矩阵 Q, 使得

$$B=PAQ.$$

引理 13.1.5 λ 矩阵的相抵关系为等价关系.

证明由直接验证可知, 在此略去.

由此立即有

定理 13.1.4 设 λ 矩阵 A 的秩为 r. 记 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 为 A 的不变因式组, 则有

$$d_1(\lambda)|d_2(\lambda), d_2(\lambda)|d_3(\lambda), \cdots, d_{r-1}(\lambda)|d_r(\lambda).$$

且 A 相抵于标准形

$$\begin{pmatrix} \begin{pmatrix} d_1(\lambda) & & \\ & \ddots & \\ & & d_r(\lambda) \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix}.$$

因此 λ 矩阵在相抵下的全系不变量为它的不变因式组．所以 λ 矩阵在相抵下的全系不变量为它的行列式因式组.

注意在上面建立的理论中，可限制在有理数域、实数域及复数域上分别讨论．而结果是完全相同的．下面限制在复数域上讨论．由代数基本定理．所以可以给出

定义 设 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 是秩为 r 的 $n\times m\lambda$ 矩阵 A 的不变因式组，则

$$\begin{aligned} d_1(\lambda) &= (\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}}\cdots(\lambda-\lambda_s)^{e_{1s}}, \\ d_2(\lambda) &= (\lambda-\lambda_1)^{e_{21}}(\lambda-\lambda_2)^{e_{22}}\cdots(\lambda-\lambda_s)^{e_{2s}}, \\ &\cdots\cdots\cdots \\ d_r(\lambda) &= (\lambda-\lambda_1)^{e_{r1}}(\lambda-\lambda_2)^{e_{r2}}\cdots(\lambda-\lambda_s)^{e_{rs}}, \end{aligned}$$

其中 e_{ij} 为非负整数，$1\leqslant i\leqslant r, 1\leqslant j\leqslant s$，且

$$e_{rj}\geqslant e_{r-1,j}\geqslant\cdots\geqslant e_{1j}\geqslant 0, j=1,2,\cdots,s,$$

又 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 为不同复数.

我们称多项式集

$$\{(\lambda-\lambda_j)^{e_{ij}}, e_{ij}\neq 0, 1\leqslant j\leqslant s, 1\leqslant i\leqslant r\}$$

为 λ 矩阵 A 的**初等因式组**，每个多项式 $(\lambda-\lambda_j)^{e_{ij}}$，$e_{ij}\neq 0$ 称为 λ 矩阵 A 的**初等因式**.

定理 13.1.5 $n\times m\lambda$ 矩阵的秩及初等因式组完全决定了它在相抵下的标准形，所以它们是相抵下的全系不变量.

证 由于不变因式组为 $n\times m$ λ 矩阵 A 在相抵下的全系不变量，所以只要证明可以用 A 的秩 r 及初等因式组 $\{(\lambda-\lambda_j)^{m_{ij}}, i=1,2,\cdots,n_j, j=1,2,\cdots,s, \lambda_1,\cdots, \lambda_s$ 两两不等$\}$来决定 A 的不变因

式组即可．已知A的不变因式组为$d_1(\lambda),\cdots,d_r(\lambda)$．由于$d_1(\lambda)|d_2(\lambda),\cdots,d_{r-1}(\lambda)|d_r(\lambda)$．所以将$m_{1j}, m_{2j}, \cdots, m_{n_j j}$按顺序排为$e_{rj}\geqslant e_{r-1,j}\geqslant\cdots\geqslant e_{r-n_j+1,j}>0$．记$e_{r-n_j,j}=\cdots=e_{1j}=0, j=1,2,\cdots,s$．于是有

$$e_{rj}\geqslant e_{r-1,j}\geqslant\cdots\geqslant e_{1j}\geqslant 0, j=1,2,\cdots,s.$$

因此

$$d_r(\lambda)=(\lambda-\lambda_1)^{e_{r1}}(\lambda-\lambda_2)^{e_{r2}}\cdots(\lambda-\lambda_s)^{e_{rs}},$$
$$d_{r-1}(\lambda)=(\lambda-\lambda_1)^{e_{r-1,1}}(\lambda-\lambda_2)^{e_{r-1,2}}\cdots(\lambda-\lambda_s)^{e_{r-1,s}},$$
$$\cdots\cdots\cdots$$
$$d_1(\lambda)=(\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}}\cdots(\lambda-\lambda_s)^{e_{1s}}.$$

所以$d_1(\lambda),\cdots,d_r(\lambda)$为$A$的不变因式组．显然它由初等因式组唯一决定．证完．

下面给出计算初等因式组的方法．

定理 13.1.6 设$n\times m\lambda$矩阵

$$A=\begin{pmatrix}\begin{pmatrix}f_1(\lambda) & & \\ & \ddots & \\ & & f_r(\lambda)\end{pmatrix} & 0\\ 0 & 0\end{pmatrix}$$

的秩为r，其中$f_1(\lambda),\cdots,f_r(\lambda)$为非零多项式．设

$f_i(\lambda)=a_i(\lambda-\lambda_1)^{f_{i1}}(\lambda-\lambda_2)^{f_{i2}}\cdots(\lambda-\lambda_s)^{f_{is}}, i=1,2,\cdots,r$，其中$a_i$为$f_i(\lambda)$之首项系数，$\lambda_1,\lambda_2,\cdots,\lambda_s$为$s$个不同复数，$f_{ij}$为非负整数，$i=1,2,\cdots,r, j=1,2,\cdots,s$．则

$$\{(\lambda-\lambda_j)^{f_{ij}}, f_{ij}\neq 0, 1\leqslant i\leqslant r, 1\leqslant j\leqslant s\}$$

为A的初等因式组．

证 我们来计算$n\times m\lambda$矩阵A的行列式因式．由定义可知

$$D_1(\lambda)=(f_1(\lambda),\cdots,f_r(\lambda))=\prod_{j=1}^{s}(\lambda-\lambda_j)^{\min(f_{1j},\cdots,f_{rj})},$$

$$D_2(\lambda)=f_i(\lambda)f_l(\lambda),\ 1\leqslant i<l\leqslant r \text{ 的最大公因式}$$

$$=\prod_{j=1}^{s}(\lambda-\lambda_j)^{\min\limits_{1<i<l\leqslant r}(f_{ij}-f_{lj})}$$

于是一般有

$D_t(\lambda)=f_{i_1}(\lambda)f_{i_2}(\lambda)\cdots f_{i_t}(\lambda),\ 1\leqslant i_1<i_2<\cdots<i_t\leqslant r$ 的最大公因式，

所以有

$$D_t(\lambda)=\prod_{j=1}^{s}(\lambda-\lambda_j)^{\min\limits_{1<i_1<\cdots<i_t\leqslant r}(f_{i_1j}+f_{i_2j}+\cdots+f_{i_tj})},$$

$$t=1,2,\cdots,r.$$

将 j 固定，将 $f_{1j},f_{2j},\cdots,f_{rj}$ 按大小次序重新排为

$$0\leqslant e_{1j}\leqslant e_{2j}\leqslant\cdots\leqslant e_{rj},$$

于是

$$\min_{1\leqslant i_1<\cdots<i_t\leqslant r}(f_{i_1j}+\cdots+f_{i_tj})=e_{1j}+e_{2j}+\cdots+e_{tj},$$

取 $j=1,2,\cdots,s$，便证明了

$$D_t(\lambda)=\prod_{j=1}^{s}(\lambda-\lambda_j)^{e_{1j}+e_{2j}+\cdots+e_{tj}},\ t=1,2,\cdots,r.$$

因此 t 阶不变因式

$$d_t(\lambda)=\frac{D_t(\lambda)}{D_{t-1}(\lambda)}=\prod_{j=1}^{s}(\lambda-\lambda_j)^{e_{tj}},\ t=1,2,\cdots,r.$$

这证明了

$$\{(\lambda-\lambda_j)^{e_{tj}},\ e_{tj}\neq 0,\ t=1,2,\cdots,r,\ j=1,2,\cdots,s\}$$
$$=\{(\lambda-\lambda_j)^{f_{ij}},\ f_{ij}\neq 0,\ i=1,2,\cdots,r,\ j=1,2,\cdots,s\}$$

为 A 的初等因式组. 证完.

定理 13.1.7 设 A_1 为 $n\times m\lambda$ 矩阵，A_2 为 $p\times q\lambda$ 矩阵. 则 $(n+p)\times(m+q)\lambda$ 矩阵

$$A=\begin{pmatrix}A_1 & 0\\ 0 & A_2\end{pmatrix}$$

的全部初等因式由 A_1 的全部初等因式和 A_2 的全部初等因式合并而成.

证 由定理 13.1.4，存在可逆 λ 方阵 $P^{(n)}, Q^{(m)}, R^{(p)}, S^{(q)}$，使得

$$PA_1Q=\begin{pmatrix}\begin{pmatrix}d_1(\lambda) & & \\ & \ddots & \\ & & d_r(\lambda)\end{pmatrix} & 0\\ 0 & 0\end{pmatrix},$$

$$RA_2S=\begin{pmatrix}\begin{pmatrix}d_{r+1}(\lambda) & & \\ & \ddots & \\ & & d_{r+s}(\lambda)\end{pmatrix} & 0\\ 0 & 0\end{pmatrix},$$

其中 $r=\text{rank}(A_1), s=\text{rank}(A_2)$. 于是证明了 A 相抵于 λ 矩阵

$$\begin{pmatrix}\begin{pmatrix}d_1(\lambda) & & \\ & \ddots & \\ & & d_{r+s}(\lambda)\end{pmatrix} & 0\\ 0 & 0\end{pmatrix}.$$

由定理 13.1.6，便证明了 A 的初等因式组为 A_1 的初等因式组拼上 A_2 的初等因式组. 证完.

习题 13.1

1. 设 H_1 和 H_2 为 n 阶 Hermite 方阵. 试证：λ 方阵 $H_1+\lambda H_2$ 的不变因式为实系数多项式.

2. 给定 n 个整数 $a_1, a_2, \cdots, a_n$. 设它们的最大公因子.

$$(a_1, \cdots, a_n)=d.$$

试证：存在一个行列式等于 d 的 n 阶整数方阵，以 $a_1, \cdots, a_n$ 为指定的行或列.

8. 以整数为元素的 $n\times m$ 矩阵称为**整数矩阵**. 试仿照 λ 矩阵的理论来建立整数矩阵在相抵下的标准形理论. 即定义行列式因子，不变因子和初等因子. 定义初等变换、相抵以及求相抵下的标准形和决定整数矩阵在相抵下的全系不变量. 这里要注意的是：整数方阵可逆当且仅当它的行列式等于 ± 1.

§13.2 复方阵在相似下的标准形

在§7.2,我们已经给出了复方阵在相似下的 Jordan 标准形. 但是在实际计算 Jordan 标准形时，产生了困难. 这一节利用 λ 矩阵在相抵下的标准形理论，给出复方阵在相似下的 Jordan 标准形的具体计算方法.

引理 13.2.1 设 D 为 n 阶常数复方阵，$T(\lambda)$ 为 n 阶 λ 方阵. 则唯一存在 n 阶 λ 方阵 $T_1(\lambda)$, $T_2(\lambda)$ 及 n 阶常数复方阵 $T,\tilde{T}$ 使得

$$T(\lambda)=(\lambda I-D)T_1(\lambda)+T=T_2(\lambda)(\lambda I-D)+\tilde{T}.$$

证 由于 λ 矩阵的每个元素为 λ 的多项式，所以

$$T(\lambda)=\lambda^m T_0+\lambda^{m-1}T_1+\cdots+\lambda T_{m-1}+T_m,$$

其中 $T_0,T_1,\cdots,T_m$ 为 n 阶常数复方阵，又 $T_0\neq 0$. 于是

$$T(\lambda)=(\lambda I-D)(\lambda^{m-1}T_0)+\lambda^{m-1}(DT_0+T_1)$$
$$+\lambda^{m-2}T_2+\cdots+T_m.$$

对 m 作归纳法，便证明了存在 n 阶 λ 方阵 $T_1(\lambda)$ 及 n 阶常数复方阵 T, 使得 $T(\lambda)=(\lambda I-D)T_1(\lambda)+T$.

再证唯一性. 今若

$$T(\lambda)=(\lambda I-D)T_1(\lambda)+T=(\lambda I-D)Q_1(\lambda)+Q,$$

其中 $T_1(\lambda)$, $Q_1(\lambda)$ 都是 n 阶 λ 方阵，T,Q 都是 n 阶常数复方阵. 于是有

$$(\lambda I-D)(T_1(\lambda)-Q_1(\lambda))=Q-T.$$

设 $T_1(\lambda)=Q_1(\lambda)$, 自然有 $T=Q$. 设 $T_1(\lambda)\neq Q_1(\lambda)$, 于是

$$T_1(\lambda)-Q_1(\lambda)=\lambda^s T_0+\lambda^{s-1}T_1+\cdots+T_s,$$

其中 $T_0\neq 0, T_1, \cdots, T_s$ 为 n 阶常数复方阵. 而

$$(\lambda I-D)(\lambda^s T_0+\lambda^{s-1}T_1+\cdots+T_s)=\lambda^{s+1}T_0+\lambda^s\tilde{T}_1+\cdots+\tilde{T}_{s+1}$$
$$=Q-T.$$

其中 $\tilde{T}_1, \cdots, \tilde{T}_{s+1}$ 为 n 阶常数复方阵. 双方比较 λ 之系数, 便证明了 $T_0=0$. 这和 $T_0\neq 0$ 矛盾. 所以证明了分解唯一性.

考虑 $T(\lambda)', D'$, 则有 $T(\lambda)'=(\lambda I-D')T_2(\lambda)'+\tilde{T}'$. 双方取转置, 便证明了 $T(\lambda)=T_2(\lambda)(\lambda I-D)+\tilde{T}$. 证完.

定理 13.2.1 n 阶复方阵 A 和 B 相似当且仅当 n 阶 λ 方阵 $\lambda I-A$ 和 $\lambda I-B$ 相抵.

证 设 A 和 B 相似, 即存在 n 阶非异常数复方阵 P, 使得 $B=P^{-1}AP$. 于是有 $\lambda I-B=P^{-1}(\lambda I-A)P$. 这证明了 $\lambda I-A$ 和 $\lambda I-B$ 相抵.

反之, 若 $\lambda I-A$ 和 $\lambda I-B$ 相抵. 所以存在 n 阶可逆 λ 方阵 $P(\lambda)$ 及 $Q(\lambda)$, 使得

$$(\lambda I-B)Q(\lambda)=P(\lambda)(\lambda I-A).$$

目的是证明 A 和 B 相似. 为此, 利用引理 13.2.1, 所以有

$$P(\lambda)=(\lambda I-B)P_1(\lambda)+P,$$
$$P(\lambda)^{-1}=(\lambda I-A)R_1(\lambda)+R,$$
$$Q(\lambda)=Q_1(\lambda)(\lambda I-A)+Q.$$

于是有 $PR=I$, 即 P 非异, 且 $P^{-1}=R$. 事实上, 由

$$I=P(\lambda)P(\lambda)^{-1}=P(\lambda)[(\lambda I-A)R_1(\lambda)+R]$$
$$=P(\lambda)(\lambda I-A)R_1(\lambda)+P(\lambda)R.$$

由于 $P(\lambda)(\lambda I-A)=(\lambda I-B)Q(\lambda)$, 代入有

$$I=(\lambda I-B)Q(\lambda)R_1(\lambda)+(\lambda I-B)P_1(\lambda)R+PR$$
$$=(\lambda I-B)T(\lambda)+PR$$

其中 $T(\lambda)=Q(\lambda)R_1(\lambda)+P_1(\lambda)R$. 我们来证 $T(\lambda)=0$. 设若不然, 则有

$$T(\lambda)=\lambda^t T_0+\lambda^{t-1}T_1+\cdots+T_t,$$

其中 $T_0\neq 0, T_1, \cdots, T_t$ 为 n 阶常数复方阵. 代入,有

$$I=(\lambda I-B)(\lambda^t T_0+\lambda^{t-1}T_1+\cdots+T_t)+PR$$
$$=\lambda^{t+1}T_0+\lambda^t\tilde{T}_1+\cdots+\tilde{T}_{t+1}+PR,$$

其中 $\tilde{T}_1, \cdots, \tilde{T}_{t+1}$ 为 n 阶常数复方阵. 比较 λ^{t+1} 之系数, 有 $T_0=0$. 这和 $T_0\neq 0$ 矛盾. 所以证明了 $T(\lambda)=0$. 因此 $PR=I$, 即 $\det P\neq 0$, 且 $R=P^{-1}$.

现在来证明 $B=PAP^{-1}$. 事实上,由

$$(\lambda I-B)Q(\lambda)=P(\lambda)(\lambda I-A),$$

有

$$(\lambda I-B)[Q_1(\lambda)(\lambda I-A)+Q]=[(\lambda I-B)P_1(\lambda)+P](\lambda I-A)$$

即有

$$(\lambda I-B)(P_1(\lambda)-Q_1(\lambda))(\lambda I-A)=(\lambda I-B)Q-P(\lambda I-A).$$

我们来证 $P_1(\lambda)-Q_1(\lambda)=0$. 设若不然,则

$$P_1(\lambda)-Q_1(\lambda)=\lambda^t T_0+\lambda^{t-1}T_1+\cdots+T_t,$$

其中 $T_0\neq 0, T_1, \cdots, T_t$ 为 n 阶常数方阵. 代回原式,所以等式左边关于 λ 的最高次项为 $\lambda^{t+2}T_0$, 等式右边关于 λ 的最高次项为 $\lambda(Q-P)$. 由 $T_0\neq 0$ 便导出矛盾. 所以证明了 $P_1(\lambda)-Q_1(\lambda)=0$, 因此有 $\lambda(Q-P)+PA-BQ=0$. 所以 $Q=P$, $PA=BQ$. 这证明了

$$A=P^{-1}BQ=P^{-1}BP.$$

至此证明了 A 和 B 相似. 定理证完.

由定理 13.2.1 和 §13.1,所以有

定理13.2.2 n 阶复方阵 A 和 B 相似当且仅当 n 阶 λ 方阵 $\lambda I-A$ 和 $\lambda I-B$ 的初等因式组相同.

且我们可以自然地引进

定义 给定 n 阶复方阵 A. 则 n 阶 λ 方阵 $\lambda I-A$ 的行列式因式、不变因式、初等因式分别称为 n 阶复方阵 A 的**行列式因式、不**

变因式、初等因式.

于是求 n 阶复方阵 A 在相似下的标准形的问题可以变为先计算 $\lambda I-A$ 的行列式因式，再算出不变因式，从而算出初等因式组. 或者用初等变换将 n 阶 λ 方阵 $\lambda I-A$ 化为对角形，再将对角元素分解为一次因式之乘积，从而定出了 $\lambda I-A$ 的初等因式组. 然后，我们设法造出一个 n 阶复方阵 J，使得它的初等因子组和给定的初等因子组相同. 由定理 13.2.2，便证明了它们相似. 当然，我们并没有给出好办法来计算非异复方阵 P，使得 $J=P^{-1}AP$.

在下面我们将上面描述的过程，严格地表达如下：

引理 13.2.2 给定复数 λ_0 及自然数 e_0，则 n_0 阶复方阵

$$J_0=\begin{pmatrix}\lambda_0 & 1 & & \\ & & \ddots & \\ & \ddots & & 1 \\ & & & \lambda_0\end{pmatrix}$$

以 $\{(\lambda-\lambda_0)^{e_0}\}$ 为初等因式组. 所以我们称 J_0 为由初等因式 $(\lambda-\lambda_1)^{e_0}$ 决定的 **Jordan 块.**

证 令 e_0 阶 λ 矩阵

$$\lambda I-J_0=\begin{pmatrix}\lambda-\lambda_0 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & \lambda-\lambda_0\end{pmatrix}.$$

它的行列式因式分别为

$$D_1(\lambda)=1, D_2(\lambda)=1,\cdots,\quad D_{e_0-1}(\lambda)=1,\quad D_{e_0}(\lambda)=(\lambda-\lambda_0)^{e_0}$$

所以它的不变因式分别为

$$d_1(\lambda)=1,\ d_2(\lambda)=1,\cdots,\ d_{e_0-1}(\lambda)=1,\ d_{e_0-1}(\lambda)=(\lambda-\lambda_0)^{e_0}.$$

因此 $\lambda I-J_0$ 的初等因式组为 $\{(\lambda-\lambda_0)^{e_0}\}$. 证完.

定理 13.2.3 设 n 阶复方阵 A 的初等因式组为

$$\{(\lambda-\lambda_1)^{e_1},\quad (\lambda-\lambda_2)^{e_2},\cdots,(\lambda-\lambda_t)^{e_t}\}$$

其中 $e_1,\cdots,e_t$ 为自然数，$\lambda_1,\lambda_2,\cdots,\lambda_t$ 为复数，它们可能有相等的．则有

$$e_1+e_2+\cdots+e_t=n,$$

且这时 A 必相似于准对角形

$$J=\operatorname{diag}(J_1^{(e_1)},\cdots,J_t^{(e_t)}),$$

称为 A 的 **Jordan 标准形**，其中

$$\mathbf{J}_i^{(e_i)}=\begin{pmatrix}\lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{pmatrix}^{(e_i)},\quad i=1,2,\cdots,t$$

为由初等因式 $(\lambda-\lambda_i)^{e_i}$ 决定的 Jordan 块．且不计 Jordan 块的排列次序，Jordan 标准形唯一．

证　由于 $\det(\lambda I-A)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_n$ 为非零多项式，所以 $\lambda I-A$ 的秩为 n．记 $d_1(\lambda),\cdots,d_n(\lambda)$ 为它的 n 个不变因式，则有 $\det(\lambda I-A)=c_0d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda)$，其中 c_0 为非零复数．但是 $d_1(\lambda),\cdots,d_n(\lambda)$ 的首项系数都等于 1，所以证明了 $c_0=1$．由初等因式组的定义可知

$$\det(\lambda I-A)=(\lambda-\lambda_1)^{e_1}(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_t)^{e_t}.$$

双方比较次数，便证明了

$$e_1+e_2+\cdots+e_t=n.$$

由定理 13.1.7 可知

$$\lambda I^{(n)}-\mathbf{J}=\operatorname{diag}(\lambda I^{(e_1)}-J_1^{(e_1)},\cdots,\lambda I^{(e_t)}-J_t^{(e_t)})$$

的初等因式组为 $\lambda I^{(e_1)}-J_1^{(e_1)},\cdots,\lambda I^{(e_t)}-J_t^{(e_t)}$ 的初等因式组之并集．由引理 13.2.2 可知它们是 $\{(\lambda-\lambda_1)^{e_1},\cdots,(\lambda-\lambda_t)^{e_t}\}$．这证明了 A 和 J 的初等因式组相同．由定理 13.2.2，所以 n 阶复方阵 A 和 Jordan 标准形 J 相似．

又将 Jordan 标准形 J 的准对角块次序改变，但是初等因式组仍然相同，所以由定理 13.2.2 可知它们仍然相似. 证完.

注意 1 虽然，Jordan 标准形为上三角方阵，对角元素构成 n 阶复方阵 A 的所有特征根. 且同一特征根可能在几个Jordan 块中出现. 所以 Jordan 块的阶数不超过相应特征根的重数.

注意 2 可以有各种标准形，即由于初等因式 $(\lambda-\lambda_0)^{e_0}$ 的 Jordan 块

$$\begin{pmatrix} \lambda_0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{pmatrix}^{(e_0)}$$

相似于下面的简单方阵:

$$\begin{pmatrix} \lambda_0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & \ddots & \lambda_0 \\ & & & 1 \end{pmatrix}^{(e_0)}$$

所以我们也可以取这种下三角方阵为Jordan 块. 在 $\lambda_0\neq 0$时，它还可以相似于下面的简单方阵:

$$\lambda_0\begin{pmatrix} 1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix}^{(e_0)}$$

这只要计算 λ 方阵

$$\lambda I^{(e_0)}-\lambda_0\begin{pmatrix} 1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix}^{(e_0)}=\begin{pmatrix} \lambda-\lambda_0 & -\lambda_0 & & \\ & \ddots & \ddots & \\ & & & -\lambda_0 \\ & & & \lambda-\lambda_0 \end{pmatrix}$$

的行列式因式，便能证明这个断言.

下面利用定理 13.2.1 的证明来给出 n 阶非异复方阵 P 的一种计算方法.

定理13.2.4 给定 n 阶复方阵 A，设 J 为 A 的 Jordan 标准形，即存在一系列初等变换，可将 λ 方阵 $\lambda I-A$ 变为 λ 方阵 $\lambda I-J$. 换句话说，存在 n 阶可逆 λ 方阵 $P(\lambda)$ 及 $Q(\lambda)$，使得

$$P(\lambda)(\lambda I-A)Q(\lambda)=\lambda I-J.$$

记

$$P(\lambda)=\lambda^t P_0+\lambda^{t-1}P_1+\cdots+P_t,$$

其中 $P_0\neq 0, P_1,\cdots,P_t$ 为 n 阶常数复方阵. 则

$$P=J^t P_0+J^{t-1}P_1+\cdots+P_t$$

有 $\det P\neq 0$，且

$$PAP^{-1}=J.$$

证 今存在 n 阶 λ 方阵 $P_1(\lambda)$ 及 n 阶常数复方阵 P，使得

$$P(\lambda)=(\lambda I-J)P_1(\lambda)+P.$$

由定理 13.2.1 之证明可知 $\det P\neq 0$，且 $PAP^{-1}=J$. 所以问题化为给出 P 由 J 决定的明显公式. 为此，记

$$P_1(\lambda)=\lambda^{t-1}D_0+\lambda^{t-2}D_1+\cdots+D_{t-1},$$

其中 $D_0\neq 0, D_1,\cdots,D_{t-1}$ 为 n 阶常数复方阵. 所以由

$$\begin{aligned}P(\lambda)&=\lambda^t P_0+\lambda^{t-1}P_1+\cdots+P_t=(\lambda I-J)P_1(\lambda)+P\\&=(\lambda I-J)(\lambda^{t-1}D_0+\lambda^{t-2}D_1+\cdots+D_{t-1})+P\\&=\lambda^t D_0+\lambda^{t-1}(D_1-JD_0)+\cdots+\lambda(D_{t-1}-JD_{t-2})+P-JD_{t-1}.\end{aligned}$$

双方比较同类项系数，则有

$$D_0=P_0$$

$$D_1-JD_0=P_1,$$

$$\cdots\cdots$$

$$D_{t-1}-JD_{t-2}=P_{t-1},$$

$$P-JD_{t-1}=P_t$$

依次消去 $D_0, D_1, \cdots, D_{t-1}$，便有

$$\begin{aligned}P &= P_t + JD_{t-1} \\ &= P_t + J(P_{t-1}+JD_{t-2}) = P_t + JP_{t-1} + J^2 D_{t-2} \\ &= \cdots \\ &= P_t + JP_{t-1} + J^2 P_{t-2} + \cdots + J^t P_0.\end{aligned}$$

证完.

Jordan 标准形有很多应用. 下面给出几个应用.

定理13.2.5 n阶复方阵 A 的极小多项式为 λ 方阵 $\lambda I - A$ 的第n个不变因式 $d_n(\lambda)$.

证 由于在相似下极小多项式不改变，所以我们只要证明 Jordan 标准形 J 的极小多项式为 λ 方阵 $\lambda I - J$ 的第 n 个不变因式 $d_n(\lambda)$就行了. 设

$$d_n(\lambda) = (\lambda-\lambda_1)^{f_1}(\lambda-\lambda_2)^{f_2}\cdots(\lambda-\lambda_s)^{f_s},$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda s$ 为不同复数，$f_1, f_2, \cdots, f_s$ 为自然数，且 Jordan 标准形的任一初等因式必形如$(\lambda-\lambda_j)^{l_j}$，其中$1 \leqslant l_j \leqslant f_j$. 由它决定的Jordan 块为

$$\begin{pmatrix} \lambda_j & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda_j \end{pmatrix}^{(l_j)},$$

因此

$$\left[\begin{pmatrix} \lambda_j & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda_j \end{pmatrix}^{(l_j)} - \lambda_j I^{(l_j)}\right]^{(f_j)} = \left[\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{pmatrix}^{(l_j)}\right]^{(l_j)} = 0.$$

所以

$$d_n\left(\begin{pmatrix}\lambda_j & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda_j\end{pmatrix}^{(l_j)}\right)=0.$$

显然，$d_n(\mathrm{diag}(J_1,\cdots,J_s))=\mathrm{diag}(d_n(J_1),\cdots,d_n(J_s))$. 这证明了 $d_n(J)=0$. 记 J 的极小多项式为 $m(\lambda)$，则有 $m(\lambda)\mid d_n(\lambda)$，所以

$$m(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s},$$

其中 $m_1,\cdots,m_s$ 为非负整数，且 $0\leqslant m_j\leqslant f_j, 1\leqslant j\leqslant s$. 由极小多项式的定义，有 $m(J)=(J-\lambda_1 I)^{m_1}(J-\lambda_2 I)^{m_2}\cdots(J-\lambda_s I)^{m_s}=0$. 所以由初等因式 $(\lambda-\lambda_j)^{f_j}$ 决定的 Jordan 块

$$J_j=\begin{pmatrix}\lambda_j & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda_j\end{pmatrix}^{(f_j)}$$

有 $m(J_j)=0, j=1,2,\cdots,s$. 即有

$$0=m(J_j)=(J_j-\lambda_1 I^{(f_j)})^{m_1}(J_j-\lambda_2 I^{(f_j)})^{m_2}\cdots(J_j-\lambda_s I^{(f_j)})^{m_s}=$$

$$\begin{pmatrix}\lambda_j-\lambda_1 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda_j-\lambda_1\end{pmatrix}^{m_1}\begin{pmatrix}\lambda_j-\lambda_2 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda_j-\lambda_2\end{pmatrix}^{m_2}\cdots\begin{pmatrix}\lambda_j-\lambda_s & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda_j-\lambda_s\end{pmatrix}^{m_s}.$$

当 $\lambda_j\neq\lambda_k$，则 $\begin{pmatrix}\lambda_j-\lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda_j-\lambda_k\end{pmatrix}$ 非异，$k=1,2,\cdots,j-1,j+1,\cdots,s$.

这证明了

$$\left(\begin{pmatrix}0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0\end{pmatrix}^{(f_j)}\right)^{m_j}=0.$$

即证明了 $f_j \leqslant m_j, j=1,2,\cdots,s$. 由于 $m(\lambda) \mid d_n(\lambda)$，所以 $f_j \geqslant m_j$，$j=1,2,\cdots,s$. 因此 $f_j=m_j, j=1,2,\cdots,n$，即有 $m(\lambda)=d_n(\lambda)$. 证完.

定理 13.2.6 任一复方阵为两个复对称方阵的乘积，且可指定其中一个为非异的.

证 给定 n 阶复方阵 A，记 J 为 A 的 Jordan 标准形. 所以存在 n 阶非异复方阵 P，使得

$$P^{-1}AP=J=\operatorname{diag}(J_1^{(e_1)},\cdots,J_t^{(e_t)}),$$

其中

$$J_k^{(e_k)}=\begin{pmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda_k \end{pmatrix}=\begin{pmatrix} & & & 1 \\ & & \cdot^{\cdot^{\cdot}} & \\ & 1 & & \\ 1 & & & \end{pmatrix}.$$

$$\begin{pmatrix} & & & \lambda_k \\ & & & 1 \\ & & \cdot^{\cdot^{\cdot}} & \\ & \lambda_k & \cdot^{\cdot^{\cdot}} & \\ \lambda_k & 1 & & \end{pmatrix}.$$

记

$$S^{(e)_k}=\begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix},$$

它为 e_k 阶对称方阵，且 $S_k^{-1}=S_k$. 又

$$S_k^{-1}J_k=\begin{pmatrix} & & & \lambda_k \\ & & \cdot^{\cdot^{\cdot}} & 1 \\ & \lambda_k & \cdot^{\cdot^{\cdot}} & \\ \lambda_k & 1 & & \end{pmatrix}$$

仍为 e_k 阶复对称方阵. 记

$$S=\operatorname{diag}(S_1^{(e_1)},\cdots,S_t^{(e_t)})$$

为 n 阶对称方阵，则 $S^{-1}J$ 仍为复对称方阵，即有

$$S^{-1}J=(S^{-1}J)'=J'S^{-1}.$$

所以有

$$S^{-1}P^{-1}AP=P'A'(P')^{-1}S^{-1}.$$

记

$$S_2=(P')^{-1}S^{-1}P^{-1}A,$$

则

$$S_2'=A'(P')^{-1}S^{-1}P^{-1}=(P')^{-1}S^{-1}P^{-1}A=S_2,$$

即 S_2 为 n 阶复对称方阵. 记

$$S_1=PSP'$$

则 S_1 为非异复对称方阵，且有 $A=S_1S_2$.

考虑 n 阶复方阵 A'，则 $A'=S_3S_4$，其中 S_3 为 n 阶非异复对称方阵，S_4 为 n 阶复对称方阵. 于是 $A=S_4S_3$. 这证明了 n 阶复方阵 A 可分解为两个 n 阶复对称方阵的乘积，且可指定其中一个为非异的. 证完.

定理 13.2.7 给定 Jordan 标准形

$$J=\operatorname{diag}(J_1^{(e_1)},\cdots,J_t^{(e_t)}),$$

其中

$$J_k^{(e_k)}=\begin{pmatrix}\lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k\end{pmatrix}^{(e_k)},\quad k=1,2,\cdots,t.$$

设 n 阶复方阵 B 和 J 可交换. 将 B 和 J 一样分块为

$$B=\begin{pmatrix}B_{11} & \cdots & B_{1t}\\ \vdots & & \vdots\\ B_{t1} & \cdots & B_{tt}\end{pmatrix},$$

则有

（1）当 $\lambda_j \neq \lambda_k$ 时，$B_{jk}=0$；

（2）当 $\lambda_j=\lambda_k, e_j \geqslant e_k$ 时，

$$B_{jk}=\begin{pmatrix} \begin{pmatrix} b_0 & b_1 & \cdots & b_{e_k-1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & b_1 \\ & & & b_0 \end{pmatrix} \\ 0 \end{pmatrix}=\begin{pmatrix} \sum_{l=0}^{e_k-1} b_l N^l \\ 0 \end{pmatrix};$$

（3）当 $\lambda_j=\lambda_k \quad e_j<e_k$ 时，

$$B_{jk}=\begin{pmatrix} 0 & \begin{pmatrix} b_0 & b_1 & \cdots & b_{e_k-1} \\ & & \ddots & \vdots \\ & & \ddots & b_1 \\ & & & b_0 \end{pmatrix} \end{pmatrix}=\begin{pmatrix} 0 & \sum_{l=0}^{e_j-1} b_l N^l \end{pmatrix},$$

其中 $b_0, b_1, \cdots$ 是任意复数，$N=\begin{pmatrix} 0 & 1 & & \\ & & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$.

证 由 $JB=BJ$，所以有 $J_j B_{jk}=B_{jk} J_k$，$1 \leqslant j, k \leqslant t$. 记

$$B_{jk}=\begin{pmatrix} a_{11} & \cdots & a_{1e_k} \\ \vdots & & \vdots \\ a_{e_j 1} & \cdots & a_{e_j e_k} \end{pmatrix}$$

由 $J_j B_{jk}=B_{jk} J_k$，有

$$\begin{pmatrix} \lambda_j a_{11}+a_{21} & \cdots & \lambda_j a_{1e_k}+a_{2e_k} \\ \vdots & & \vdots \\ \lambda_j a_{e_j-1,1}+a_{e_j 1} & \cdots & \lambda_j a_{e_j-1,e_k}+a_{e_j e_k} \\ \lambda_j a_{e_j 1} & \cdots & \lambda_j a_{e_j e_k} \end{pmatrix}$$

$$=\begin{pmatrix} \lambda_k a_{11} & a_{11}+\lambda_k a_{12} & \cdots & a_{1,e_k-1}+\lambda_k a_{1e_k} \\ \vdots & \vdots & & \vdots \\ \lambda_k a_{e_j 1} & a_{e_j 1}+\lambda_k a_{e_j 2} & \cdots & a_{e_j,e_k-1}+\lambda_k a_{e_j e_k} \end{pmatrix}.$$

即有

$$\begin{pmatrix} (\lambda_j-\lambda_k)a_{11}+a_{21} & (\lambda_j-\lambda_k)a_{12}+a_{22}-a_{11} & \cdots & (\lambda_j-\lambda_k)a_{1e_k}+a_{2e_k}-a_{1,e_k-1} \\ \vdots & \vdots & & \vdots \\ (\lambda_j-\lambda_k)a_{e_j-1,1}+a_{e_j1} & (\lambda_j-\lambda_k)a_{e_j-1,2}+a_{e_j2}-a_{e_j-1,1} & \cdots & (\lambda_j-\lambda_k)a_{e_j-1,e_k}+a_{e_je_k}-a_{e_j-1,e_k-1} \\ (\lambda_j-\lambda_k)a_{e_j1} & (\lambda_j-\lambda_k)a_{e_j2}-a_{e_j1} & \cdots & (\lambda_j-\lambda_k)a_{e_je_k}-a_{e_j,e_k-1} \end{pmatrix}=0.$$

当$\lambda_j\neq\lambda_k$，看最后一行，则有$a_{e_j1}=0$，所以$a_{e_j2}=0,\cdots,a_{e_je_k}=0$。再从最后一行依次考察到第一行，便证明了$B_{jk}=0$.

当$\lambda_j=\lambda_k$，同样从最后一行依次考察到第一行. 于是有

$$a_{e_j1}=\cdots=a_{e_j,e_k-1}=0,\quad a_{21}=\cdots=a_{e_j1}=0,$$

$$a_{pq}=a_{p-1,q-1},\qquad p=2,\cdots,e_j,\qquad q=2,\cdots,e_k.$$

分别在$e_j\geqslant e_k$，$e_j<e_k$这两种情形来考察，便证明了定理. 证完.

在结束这一节前，我们给出实方阵在实相似下的标准形.

定义 两个n阶实方阵A和B称为**实相似的**，如果存在非异实方阵P使得$B=P^{-1}AP$.

引理 13.2.3 实方阵的实相似为等价关系.

定理13.2.8 n阶实方阵A和B实相似当且仅当它们相似.

证 由于实相似为相似，所以只要证当A和B相似，则必实相似. 今存在n阶非异复方阵$P+\sqrt{-1}Q$，其中P和Q为实方阵，使得

$$(P+\sqrt{-1}Q)^{-1}A(P+\sqrt{-1}Q)=B.$$

于是$AP+\sqrt{-1}AQ=PB+\sqrt{-1}QB$. 由于$A,B,P,Q$为实方阵，此即

$$AP=PB,\qquad AQ=QB.$$

所以任取实数 t，则有

$$A(P+tQ)=(P+tQ)B.$$

今 $P+tQ$ 为 n 阶实方阵，而

$$\det(P+tQ)=a_0t^n+a_1t^{n-1}+\cdots+a_n.$$

取 $t=\sqrt{-1}$，则有 $\det(P+\sqrt{-1}Q)\neq 0$，所以证明了 $\det(P+tQ)$ 为非零多项式，次数不超过 n. 它至多有 n 个不同实根，所以存在实数 t_0，使得 t_0 不是根，即有 $\det(P+t_0Q)\neq 0$. 所以 $P+t_0Q$ 为 n 阶非异实方阵，且有

$$B=(P+t_0Q)^{-1}A(P+t_0Q).$$

这证明了实方阵 A 和 B 实相似. 证完.

引理 13.2.4 设 $\{(\lambda-\lambda_1)^{e_1},\cdots,(\lambda-\lambda_t)^{e_t}\}$ 为 n 阶实方阵 A 的初等因式组，则 $\{(\lambda-\bar{\lambda}_1)^{e_1},\cdots,(\lambda-\bar{\lambda}_t)^{e_t}\}$ 也是 A 的初等因式组.

证 对 $\lambda I-A$ 作初等变换化为对角形 $\mathrm{diag}\,(d_1(\lambda),\cdots,d_n(\lambda))$，其中 $d_1(\lambda)\mid d_2(\lambda),\cdots,d_{n-1}(\lambda)\mid d_n(\lambda)$. 由作法可知：当 A 为实方阵，即 $\lambda I-A$ 由实系数多项式构成，则标准形也由实系数多项式构成，所以不变因式组 $\{d_1(\lambda),\cdots,d_n(\lambda)\}$ 全是实系数多项式. 这证明了初等因式组中若有 $(\lambda-\lambda_0)^{e_0}$，则必有 $(\lambda-\bar{\lambda}_0)^{e_0}$，证完.

定理 13.2.9 n 阶实方阵 A 的初等因式组为

$$\lambda^{e_1},\cdots,\lambda^{e_s},$$
$$(\lambda-\lambda_1)^{f_1},\cdots,(\lambda-\lambda_t)^{f_t},$$
$$(\lambda-\tau_1)^{g_1},\cdots,(\lambda-\tau_q)^{g_q},(\lambda-\bar{\tau}_1)^{g_1},\cdots,(\lambda-\bar{\tau}_q)^{g_q}$$

其中 $\lambda_1,\cdots,\lambda_t$ 为 A 的非零实特征根，$\tau_1,\cdots,\tau_q$ 为 A 的复特征根，其中虚部都大于零. 所以

$$\tau_j=r_je^{\sqrt{-1}\theta_j},\quad 1\leqslant j\leqslant q.$$

这时 A 有实 Jordan 标准形

$$J=\mathrm{diag}(N^{(e_1)},\cdots,N^{(e_s)},\lambda_1I^{(f_1)}+N^{(f_1)},\cdots,\lambda_tI^{(f_t)}+N^{(f_t)},$$

$$r_1L(\theta_1)^{(2g_1)},\cdots,r_qL(\theta_q)^{(2g_q)}),$$

其中

$$L(\theta)=\begin{pmatrix} M\cos\theta & M\sin\theta \\ -M\sin\theta & M\cos\theta \end{pmatrix},$$

$$N=\begin{pmatrix} 0 & 1 & & \\ & & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix},\qquad M=I+N.$$

证 由引理 13.2.4，所以实方阵 A 的初等因式组如定理所述. 因此 A 相似于 Jordan 标准形

$$\operatorname{diag}(N^{(e_1)},\cdots,N^{(e_s)},\lambda_1I^{(f_1)}+N^{(f_1)},\cdots,\lambda_tI^{(f_t)}+N^{(f_t)},$$
$$\tau_1I^{(g_1)}+N^{(g_1)},\bar\tau_1I^{(g_1)}+N^{(g_1)},\cdots,\tau_qI^{(g_q)}+N^{(g_q)},\bar\tau_qI^{(g_q)}+N^{(g_q)}).$$

由定理 13.2.8，我们只要证明 $\operatorname{diag}(\tau I^{(e)}+N^{(e)},\bar\tau I^{(e)}+N^{(e)})$ 和 $rL(\theta)^{(2e)}$ 相似就行了，其中 $\tau=re^{\sqrt{-1}\theta}$ 不是实数. 事实上，

$$\begin{pmatrix} 1 & & & \\ & \tau & & \\ & & \ddots & \\ & & & \tau^{e-1} \end{pmatrix}^{-1}\begin{pmatrix} \tau & 1 & & \\ & \tau & \ddots & \\ & & \ddots & 1 \\ & & & \tau \end{pmatrix}\begin{pmatrix} 1 & & & \\ & \tau & & \\ & & \ddots & \\ & & & \tau^{e-1} \end{pmatrix}=\tau(I^{(e)}+N^{(e)})=\tau M^{(e)},$$

$$\begin{pmatrix} 1 & & & \\ & \bar\tau & & \\ & & \ddots & \\ & & & \bar\tau^{e-1} \end{pmatrix}^{-1}\begin{pmatrix} \bar\tau & 1 & & \\ & \bar\tau & \ddots & \\ & & \ddots & 1 \\ & & & \bar\tau \end{pmatrix}\begin{pmatrix} 1 & & & \\ & \bar\tau & & \\ & & \ddots & \\ & & & \bar\tau^{e-1} \end{pmatrix}=\bar\tau M^{(e)}.$$

所以 $\operatorname{diag}(\tau I^{(e)}+N^{(e)},\bar\tau I^{(e)}+N^{(e)})$，相似于 $\operatorname{diag}(\tau M^{(e)},\bar\tau M^{(e)})$.
而

$$\left[\frac{1}{\sqrt2}\begin{pmatrix} I & -\sqrt{-1}I \\ I & \sqrt{-1}I \end{pmatrix}\right]^{-1}\begin{pmatrix} \tau M & 0 \\ 0 & \bar\tau M \end{pmatrix}\left[\frac{1}{\sqrt2}\begin{pmatrix} I & -\sqrt{-1}I \\ I & \sqrt{-1}I \end{pmatrix}\right]$$
$$=\left(\frac{1}{2}\right)\begin{pmatrix} I & I \\ \sqrt{-1}I & -\sqrt{-1}I \end{pmatrix}\begin{pmatrix} \tau M & -\sqrt{-1}\tau M \\ \bar\tau M & \sqrt{-1}\bar\tau M \end{pmatrix}$$

$$=\begin{pmatrix} (\mathrm{Re}(\tau))M & (\mathrm{Im}(\tau))M \\ -(\mathrm{Im}(\tau))M & (\mathrm{Re}(\tau))M \end{pmatrix}=\tau L(\theta)$$

这证明了 $\mathrm{diag}(\tau M,\bar{\tau}M)$ 和 $rL(\theta)$ 相似，所以 $rL(\theta)$ 和 $\mathrm{diag}(\tau I+N,\bar{\tau}I+N)$ 相似. 至此证明了定理.

习题 13.2

1. 一组两两可交换的方阵如果都相似于对角阵，则它们能同时经过一个非异方阵 P 相似于对角形.

2. 试证：存在置换方阵 P，使得

$$P\begin{pmatrix} (a_{ij}^{(11)}) & \cdots & (a_{ij}^{(1m)}) \\ \vdots & & \vdots \\ (a_{ij}^{(m1)}) & \cdots & (a_{ij}^{(mm)}) \end{pmatrix}P^{-1}=\begin{pmatrix} (a_{11}^{(pq)}) & \cdots & (a_{1n}^{(pq)}) \\ \vdots & & \vdots \\ (a_{n1}^{(pq)}) & \cdots & (a_{nn}^{(pq)}) \end{pmatrix},$$

其中

$$(a_{ij}^{(uv)})=\begin{pmatrix} a_{11}^{(uv)} & \cdots & a_{1n}^{(uv)} \\ \vdots & & \vdots \\ a_{n1}^{(uv)} & \cdots & a_{nn}^{(uv)} \end{pmatrix},\quad (a_{uv}^{(pq)})=\begin{pmatrix} a_{uv}^{(11)} & \cdots & a_{uv}^{(1m)} \\ \vdots & & \vdots \\ a_{uv}^{(m1)} & \cdots & a_{uv}^{(mm)} \end{pmatrix}.$$

利用这一事实，试证下列方阵互相相似：

(i) $$\begin{pmatrix} I^{(n)} & I^{(n)} & & & \\ & I^{(n)} & I^{(n)} & & \\ & & I^{(n)} & \ddots & \\ & & & \ddots & I^{(n)} \\ & & & & I^{(n)} \end{pmatrix}^{(mn)},\quad \begin{pmatrix} M^{(m)} & & \\ & \ddots & \\ & & M^{(m)} \end{pmatrix}^{(mn)};$$

(ii) $$\begin{pmatrix} \begin{pmatrix} 0 & M^{(m)} \\ M^{(m)} & 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} 0 & M^{(m)} \\ M^{(m)} & 0 \end{pmatrix} \end{pmatrix}^{(2mn)},\quad \begin{pmatrix} R & R & & & \\ & R & R & & \\ & & R & \ddots & \\ & & & \ddots & R \\ & & & & R \end{pmatrix}^{(2mn)},$$

其中

$$R^{(2n)}=\mathrm{diag}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\cdots,\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right);$$

(iii) 若

$$\det\begin{pmatrix} A_1^{(n)} & B_1^{(n)} \\ C_1^{(n)} & D_1^{(n)} \end{pmatrix} \neq 0,$$

试证：

$$\det\begin{pmatrix} \begin{pmatrix} A_m & \cdots & A_1 \\ \vdots & \iddots & \\ A_1 & & \end{pmatrix} & \begin{pmatrix} B_m & \cdots & B_1 \\ \vdots & \iddots & \\ B_1 & & \end{pmatrix} \\ \begin{pmatrix} C_m & \cdots & C_1 \\ \vdots & \iddots & \\ C_1 & & \end{pmatrix} & \begin{pmatrix} D_m & \cdots & D_1 \\ \vdots & \iddots & \\ D_1 & & \end{pmatrix} \end{pmatrix} \neq 0.$$

其中 $A_1, A_2, \cdots, A_m$, $B_1, B_2, \cdots$, B_m, C_1, C_2, $\cdots, C_m, D_1, D_2, \cdots, D_m$ 都是 n 阶方阵.

3. 试证：A 相似于对角形当且仅当 A 的初等因式都是一次的.

4. 试证：适合条件 $A^p=I$ 的方阵都相似于对角形.

5. 试求 N^p 的 Jordan 标准形，其中 $N^{(n)}=\begin{pmatrix} 0 & 1 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$.

6. 设 A 为 n 阶方阵，且 $\mathrm{tr}A=0$. 试证：A 相似于 n 阶方阵 B，其中 B 的对角元素都是零. 又存在 n 阶方阵 X, Y，使得

$$[X, Y]=XY-YX=A.$$

7. 设 A 为 n 阶方阵，记 $A^m=(a_{ij}^{(m)})$, $m=1,2,\cdots$. 设若存在正数 M，使得

$$|a_{ij}^{(m)}|<M, \qquad i, j=1,2,\cdots,n, m=1,2,\cdots,$$

则 A 称为幂有界方阵. 试证：A 为幂有界方阵当且仅当 A 的特征根的模 $\leqslant 1$，且模为 1 的特征根的初等因式都是 1 次的.

8. 试证：任一方阵和它的转置方阵都相似.

9. 设 A 和 B 都是 n 阶方阵，且 $\det A \neq 0$. 试证：AB 和 BA 相似.

10. 设 A 的特征根都等于 1，试证：A 和 A 的任意方幂都相似.

11. 给定 n^2 个 n 阶非零方阵 $A_{jk}, j, k=1,2,\cdots,n$. 设有

$$A_{ij}A_{pq}=\delta_{jp}A_{iq}, \qquad i, j, p, q=1,2,\cdots,n.$$

试证：存在 n 阶非异方阵 P，使得

$$P^{-1}A_{ij}P=E_{ij}, \quad i, j=1,2,\cdots,n,$$

其中 E_{ij} 为第 i 行，第 j 列元素为 1，其余元素为零的 n 阶方阵.

12. 给定自然数 p，试定出所有 n 阶方阵 A，使得 A 和 A^p 相似.

13. 设A和B分别为n阶及m阶方阵，X为nm个独立未知数构成的$n\times m$矩阵. 什么时候矩阵方程

$$AX=XB$$

有非零解X？

14. 试证:存在一个n阶非异方阵P,使得它可同时使所有轮迴方阵

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix}$$

相似于对角形.

15. 用归纳法定义$(s_1,s_2,\cdots,s_m)$轮迴方阵如下:当$m=1$即为通常的s_1阶轮迴方阵. 假设已定义好$(s_1,\cdots,s_{m-1})$轮迴方阵. 我们称A为$(s_1,s_2,\cdots,s_m)$轮迴方阵,如果

$$A=\begin{pmatrix} B_0 & B_1 & \cdots & B_{s_m-1} \\ B_{s_m-1} & B_0 & \cdots & B_{s_m-2} \\ \vdots & \vdots & & \vdots \\ B_1 & B_2 & \cdots & B_0 \end{pmatrix},$$

其中$B_0,B_1,\cdots,B_{s_m-1}$都是$(s_1,s_2,\cdots,s_{m-1})$轮回方阵. 试证:

(i) A为$n=s_1s_2\cdots s_m$阶方阵;

(ii) 存在n阶非异方阵P，使得它可同时使所有$(s_1,s_2,\cdots,s_m)$轮迴方阵相似于对角形.

§13.3 方阵函数和方阵幂级数

这一节也是Jordan标准形理论的应用，我们只考虑复数域的情形. 为了引出方阵纯函数的概念. 我们先讨论方阵多项式函数.

给定n^2个独立未知数x_{ij}, $i,j=1,2,\cdots,n$. 它组成n阶方阵

$$X=\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}.$$

给定 $m+1$ 个复常数 $a_0, a_1, \cdots, a_m$，则 n 阶方阵

$$f(X)=a_0X^m+a_1X^{m-1}+\cdots+a_{m-1}X+a_mI^{(n)}$$

称为方阵 X 的多项式，简称为**方阵多项式**.

引理 13.3.1 设 $f(X)$ 为方阵多项式.

(1) 任取 n 阶非异方阵 P，则有

$$f(P^{-1}XP)=P^{-1}f(X)P;$$

(2) 当 X 为准对角方阵 $X=\mathrm{diag}(X_1,\cdots,X_t)$ 时有

$$f(\mathrm{diag}(X_1,\cdots,X_t))=\mathrm{diag}(f(X_1),\cdots,f(X_t));$$

(3) 设 $A=\lambda_0I^{(n)}+N^{(n)}$ 为 Jordan 块，则有

$$f(A)=f(\lambda_0I+N)=\sum_{l=0}^{n-1}\frac{1}{l!}f^{(l)}(\lambda_0)N^l.$$

证 由多项式函数的定义可知性质(1)与(2)显然成立. 为了证明引理，只要证(3)，记 $f(X)=\sum_{j=0}^{m}a_jX^{m-j}$，于是

$$\begin{aligned}f(A)=f(\lambda_0I+N)&=\sum_{j=0}^{m}a_j(\lambda_0I+N)^{m-j}\\&=\sum_{j=0}^{m}a_j\sum_{l=0}^{m-j}\binom{m-j}{l}\lambda_0^{m-j-l}N^l\\&=\sum_{l=0}^{m}\left(\sum_{j=0}^{m-l}\frac{(m-j)!}{l!(m-j-l)!}a_j\lambda^{m-j-l}\right)N^l\end{aligned}$$

记 $f(x)=a_0x^m+a_1x^{m-1}+\cdots+a_m$ 为通常之多项式. 则

$$f^{(l)}(x)=\sum_{j=0}^{m-l}\frac{(m-j)!}{(m-j-l)!}a_jx^{m-l-j}.$$

因此证明了

$$f(A)=\sum_{l=0}^{\infty}\frac{1}{l!}f^{(l)}(\lambda_0)N^l$$

证完

由引理 13.3.1，有下面自然的推论.

推论 给定 n 阶复方阵 A，记 $J=\mathrm{diag}(J_1^{(e_1)}, \cdots, J_t^{(e_t)})$ 为 A 的 Jordan 标准形，其中

$$J_k^{(e_k)}=\lambda_k I^{(e_k)}+N_k^{(e_k)}, k=1,2,\cdots,t$$

为Jordan 块，这时存在 n 阶非异复方阵 P 使得 $A=PJP^{-1}$，且对 m 次多项式 $f(x)$，则

$$f(A)=P\mathrm{diag}\left(\sum_{l=0}^{e_1-1}\frac{1}{l!}f^{(l)}(\lambda_1)N_1^l,\cdots,\sum_{l=0}^{e_t-1}\frac{1}{l!}f^{(l)}(\lambda_t)N_t^l\right)P^{-1}.$$

所以 $f(A)$ 之值只和 P 及 $f(x)$ 的各阶导数在特征根上的值有关.

在分析课中我们知道，作为多项式的自然推广为幂级数. 所以，我们可以同样讨论方阵幂级数. 为此先给出极限的定义.

定义 给定 $n\times m$ 矩阵序列 $A_p=(a_{jk}^{(p)})$. 若 nm 个序列 $\{a_{jk}^{(p)}\}$ $j=1,\cdots,n, k=1,\cdots,m$ 都收敛，且设

$$\lim_{p\to\infty}a_{jk}^{(p)}=a_{jk}, j=1,2,\cdots,n, k=1,2,\cdots,m,$$

则称 $n\times m$ 矩阵序列 $\{A_p\}$ **收敛**，且**极限**为 $n\times m$ 矩阵

$$A=(a_{jk}).$$

这时记作 $\lim_{p\to\infty}A_p=A$.

定义 给定 $n\times m$ 矩阵序列

$$B_p=\sum_{l=0}^{p}a_lA_l, \quad p=1,2,\cdots,$$

其中 $A_0, A_1,\cdots$ 为 $n\times m$ 矩阵，$a_0, a_1,\cdots$ 为复数. 若 $\{B_p\}$ 收敛于极限 B，则记作

$$B=\sum_{l=0}^{\infty}a_lA_l$$

它称为**无穷级数**、这时也称级数 $\sum_{l=0}^{\infty}a_lA_l$ **收敛于** B. 否则，称为**发散的**.

引理 13.3.2 记 $A_l=(a_{ij}^{(l)})$ 为 $n\times m$ 矩阵，$l=0,1,2,\cdots$，a_0，$a_1,\cdots$为复数. 则矩阵级数

$$\sum_{l=0}^{\infty}a_l A_l$$

收敛当且仅当 nm 个复数级数

$$\sum_{l=0}^{\infty}a_l a_{ij}^{(l)},\ i=1,2,\cdots,n,\ \ j=1,2,\cdots,m$$

都收敛.

定义 给定 n 阶复方阵 A，给定复数序列$\{a_l\}$. 则方阵级数

$$\sum_{l=0}^{\infty}a_l A^l$$

称为方阵 A 的**幂级数**. 给定 n^2 个独立未知数构成的 n 阶方阵X，则方阵函数

$$f(X)=\sum_{l=0}^{\infty}a_l X^l$$

称为**方阵幂级数**.

定理 13.3.1 给定一个复变数 z 的幂级数

$$f(z)=\sum_{l=0}^{\infty}a_l z^l.$$

记它的收敛半径为 ρ，则方阵 A 的幂级数

$$f(A)=\sum_{l=0}^{\infty}a_l A^l$$

有如下性质：记 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 A 的特征根，且

$$\lambda_0=\max_{1\leqslant j\leqslant n}|\lambda_j|,$$

则有

(i) 当$\rho>\lambda_0$时，方阵A的幂级数$f(A)=\sum_{l=0}^{\infty}a_l A^l$ 收敛，且对e。

阶 Jordan 块 $\lambda_0 I^{(e_0)}+N^{(e_0)}$，有

$$f(\lambda_0 I^{(e_0)}+N^{(e_0)})=\sum_{l=0}^{e_0-1}\frac{1}{l!}f^{(l)}(\lambda_0)N^l.$$

(ii) 当$\rho<\lambda_0$时，方阵A的幂级数 $f(A)=\sum_{l=0}^{\infty}a_l A^l$ 发散.

(iii) 当 $\rho=\lambda_0$ 时，方阵 A 的幂级数 $f(A)=\sum_{l=0}^{\infty}a_l A^l$ 收敛当且仅当对每个绝对值为ρ的特征根 λ_j，记λ_j的初等因式中最高次数为n_j，则 n_j 个数值幂级数

$$f^{(l)}(\lambda_j)=\sum_{k=l}^{\infty}\frac{k!}{(k-l)!}a_k\lambda_j^{k-l},\ l=0,1,\cdots,n_j-1$$

都收敛.

证 给定n阶复方阵 A，则存在n阶非异复方阵P使得$P^{-1}AP=J$ 为Jordan标准形，其中 $J=\mathrm{diag}(J_1^{(e_1)},\cdots,J_t^{(e_t)})$，$J_k^{(e_k)}$为由初等因式$(\lambda-\lambda_k)^{e_k}$ 决定的 Jordan 块. 由幂级数定义可知

$$f(A)=\sum_{l=0}^{\infty}a_l A^l=\sum_{l=0}^{\infty}a_l(PJP^{-1})^l=P\left(\sum_{l=0}^{\infty}a_l J^l\right)P^{-1}=Pf(J)P^{-1},$$

而

$$\begin{aligned}f(J)&=\sum_{l=0}^{\infty}a_l(\mathrm{diag}(J_1,\cdots,J_t))^l\\&=\mathrm{diag}\left(\sum_{l=0}^{\infty}a_l J_1^l,\cdots,\sum_{l=0}^{\infty}a_l J_t^l\right)\\&=\mathrm{diag}(f(J_1),\cdots,f(J_t)).\end{aligned}$$

又记 $J_0=\lambda_0 I^{(e_0)}+N^{(e_0)}$，则

$$\begin{aligned}f(J_0)&=\sum_{l=0}^{\infty}a_l(\lambda_0 I+N)^l=\sum_{l=0}^{\infty}a_l\sum_{k=0}^{l}\binom{l}{k}\lambda_0^{l-k}N^k\\&=\sum_{k=0}^{\infty}\sum_{l=k}^{\infty}\frac{l!}{k!(l-k)!}a_l\lambda_0^{l-k}N^k=\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(\lambda_0)N^k.\end{aligned}$$

由于 N 为 e_0 阶方阵，且 $N^{e_0}=0$. 所以

$$f(J_0)=\sum_{k=0}^{e_0-1}\frac{1}{k!}f^{(k)}(\lambda_0)N^k.$$

由此证明了当 $f(z),f'(z),\cdots,f^{(e_i-1)}(z)$ 在 $z=\lambda_i$ 时收敛，$i=1,2,\cdots,t$，则 $f(A)$ 收敛. 由通常幂级数理论便证明了定理. 证完.

利用这个定理，我们可以引进一系列初等方阵函数. 它们是

(a) $(I+A)^\alpha=\sum_{k=0}^{\infty}\binom{\alpha}{k}A^k$,

其中 α 为正实数，且

$$\binom{\alpha}{k}=\frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}.$$

(b) $\sin A=\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(2k-1)!}A^{2k-1}$.

(c) $\cos A=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}A^{2k}$.

(d) $e^A=\exp A=\sum_{k=0}^{\infty}\frac{1}{k!}A^k$.

(e) $\log(I+A)=\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}A^k$.

定义 给定通常复幂级数 $f(z)=\sum_{l=0}^{\infty}a_lz^l$，则它的收敛半径 ρ 称为方阵幂级数 $f(A)=\sum_{l=0}^{\infty}a_lA^l$ 的收敛半径.

由此定义可知，对方阵幂级数 $\sin A,\cos A,\exp A$，它们的收敛半径都是无穷大. 所以对任意复方阵 A，这些方阵幂级数都收敛，即有定义. 对方阵幂级数 $(I+A)^\alpha$ 及 $\log(I+A)$，它们的收敛半径为 1，所以只要方阵 A 的特征根的模小于 1，则方阵幂级数 $(I+A)^\alpha$

及$\log(I+A)$都收敛，即有定义.

特别重要的方阵幂级数为 $\exp A$.

由于矩阵之乘法不适合交换律，所以方阵幂级数也有一些和普通幂级数不同的性质.

定理 13.3.2 若 n 阶复方阵 A 和 B 可交换，则有

$$\exp A\exp B=\exp B\exp A=\exp(A+B).$$

证 今由 $\exp(A+B)=\exp(B+A)$，所以只要证$\exp A\exp B=\exp(A+B)$就够了. 今

$$\exp A\exp B=\sum_{k=0}^{\infty}\frac{1}{k!}A^k\sum_{l=0}^{\infty}\frac{1}{l!}B^l=\sum_{k,l=0}^{\infty}\frac{1}{k!l!}A^kB^l.$$

$$\exp(A+B)=\sum_{q=0}^{\infty}\frac{1}{q!}(A+B)^q=\sum_{q=0}^{\infty}\frac{1}{q!}\sum_{k=0}^{q}\binom{q}{k}A^kB^{q-k}$$

$$=\sum_{q=0}^{\infty}\sum_{k=0}^{q}\frac{1}{k!(q-k)!}A^kB^{q-k}.$$

这里用到了 $AB=BA$ 的条件. 于是

$$\exp(A+B)=\sum_{q=0}^{\infty}\sum_{k+l=q}\frac{1}{k!l!}A^kB^l.$$

这证明了

$$\exp(A+B)-\exp A\exp B=\sum_{q=0}^{\infty}\sum_{k+l=q}\frac{1}{k!l!}A^kB^l-\sum_{k,l=0}^{\infty}\frac{1}{k!l!}A^kB^l.$$

为了证明它等于零，问题化为证明序列

$$C_p=\sum_{q=0}^{p}\sum_{k+l=q}\frac{1}{k!l!}A^kB^l-\sum_{k,l=0}^{p}\frac{1}{k!l!}A^kB^l=-\sum_{\substack{k+l>p\\0\leqslant k,l\leqslant p}}\frac{-1}{k!l!}A^kB^l$$

当 $p\to\infty$ 时极限为零.

记 $A=(a_{ij}),B=(b_{ij})$,

$$d=1+\max_{1\leqslant i,j\leqslant n}(|a_{ij}|,|b_{ij}|).$$

记 $A^k=(a_{ij}^{(k)}),B^l=(b_{ij}^{(l)})$. 由归纳法不难证明.

$$|a_{ij}^{(k)}|\leqslant(nd)^k,|b_{ij}^{(l)}|\leqslant(nd)^l,1\leqslant i,j\leqslant n.$$

今 C_p 之第 i 行第 j 列元素

$$c_{ij}^{(p)}=\sum_{0\leqslant k,l\leqslant p}\ \sum_{k+l>p}\frac{1}{k!l!}\sum_{r=1}^{n}a_{ir}^{(k)}b_{rj}^{(l)}.$$

所以

$$|c_{ij}^{(p)}|\leqslant\sum_{0\leqslant k,l\leqslant p}\ \sum_{k+l>p}\frac{n}{k!l!}(nd)^{k+l}$$

$$=-n\sum_{q=0}^{p}\ \sum_{k+l=q}\frac{1}{k!l!}(nd)^{k+l}+n\sum_{k,l=0}^{p}\frac{1}{k!l!}(nd)^{k+l}$$

当$p\to\infty$,则极限小于等于$-n\exp(2nd)+n\exp(2nd)=0$. 这证明了 $\lim\limits_{p\to\infty}C_p=0$. 所以证明了 $\exp(A+B)=\exp A\exp B$. 证完.

定理 13.3.3 n 阶复方阵A非异当且仅当存在 n 阶复方阵B,使得 $A=\exp B$. 即当 n 阶复方阵 A 非异时,矩阵方程

$$\exp X=A$$

有解 B.

证 若存在方阵 B,使得 $A=\exp B$. 记 B 的 n 个特征根为 λ_1, $\lambda_2,\cdots,\lambda_n$, 于是存在 n 阶非异复方阵 P,使得

$$B=P\begin{pmatrix}\lambda_1 & & \\ & \ddots & * \\ & & \lambda_n\end{pmatrix}P^{-1}.$$

于是

$$\exp B=\sum_{l=0}^{\infty}\frac{1}{l!}B^l=P\left[\sum_{l=0}^{\infty}\frac{1}{l!}\begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ & & \lambda_n\end{pmatrix}^l\right]P^{-1}$$

$$=P\begin{pmatrix} \sum_{l=0}^{\infty}\frac{1}{l!}\lambda_1^l & & * \\ & \ddots & \\ & & \sum_{l=0}^{\infty}\frac{1}{l!}\lambda_n^l \end{pmatrix}P^{-1}$$

所以 $\exp B$ 的 n 个特征根为 $e^{\lambda_1},\cdots,e^{\lambda_n}$. 因此

$$\det\ \exp B=\prod_{j=1}^{n}e^{\lambda_j}=e^{\sum_{j=1}^{n}\lambda_j}=e^{\mathrm{tr}B}.$$

这实际上证明了一个有用的公式

$$\det(\exp B)=\exp(\mathrm{tr}B).$$

于是由 $\exp(\mathrm{tr}B)\neq 0$ 可知 $\det(\exp B)\neq 0$, 即 $\exp B=A$ 非异. *

反之, 我们来证明: 若 A 非异, 则存在方阵 B 使得 $\exp B=A$. 今对非异方阵 A, 已知存在 n 阶非异复方阵 P, 使得

$$P^{-1}AP=\mathrm{diag}(\lambda_1 M^{(e_1)},\cdots,\lambda_t M^{(e_t)}),$$

其中 $\lambda_1,\cdots,\lambda_t$ 为 A 的特征根, 它们都不等于零, 又

$$M^{(e_i)}=I^{(e_i)}+N^{(e_i)}.$$

今 $\lambda_j\neq 0$, 于是有 $\mu_j=\ln\lambda_j,\ j=1,2,\cdots,n$, 即

$$\lambda_j=e^{\mu_j},\ j=1,2,\cdots,n.$$

若对 e_0 阶方阵 M, 存在 e_0 阶方阵 D 使得 $\exp D=M$. 因此存在 e_j 阶方阵 B_j 使得 $\exp B_j=M^{(e_j)},\ j=1,2,\cdots,t$. 于是

$$\begin{aligned}A&=P\mathrm{diag}(\lambda_1 M^{(e_1)},\cdots,\lambda_t M^{(e_t)})P^{-1}\\&=P\mathrm{diag}(\exp(\mu_1 I+B_1),\cdots,\exp(\mu_t I+B_t))P^{-1}\\&=\exp[P\mathrm{diag}(\mu_1 I+B_1,\cdots,\mu_t I+B_t)P^{-1}].\end{aligned}$$

所以 $A=\exp X$ 有解 $X=P\mathrm{diag}(\mu_1 I+B_1,\cdots,\mu_t I+B_t)P^{-1}$. 因此问题化为求矩阵方程

$$\exp X=M=I+N$$

之解，其中M为 e_0 阶方阵.

今对通常之级数，有

$$\exp(\log(1+z))=1+z,\quad \forall|z|<1.$$

因此
$$\log M=\log(1+N)=\sum_{k=1}^{e_0-1}\frac{(-1)^{k-1}}{k}N^k$$

有意义. 所以利用 $N^{e_0}=0$，及

$$\sum_{l=0}^{\infty}\frac{1}{l!}\left(\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}z^k\right)^l=1+z$$

用N代入，易证

$$\sum_{l=0}^{\infty}\frac{1}{l!}\left(\sum_{k=1}^{e_0-1}\frac{(-1)^{k-1}}{k}N^k\right)^l=I+N.$$

这证明了

$$\exp(\log M)=M$$

所以矩阵方程 $\exp X=M$ 有解$\log M$. 证完.

注意：设 A为 n 阶非异复方阵，上面证明了矩阵方程$\exp X=A$ 有解. 显然它解不唯一. 事实上，由于

$$e^{2k\pi\sqrt{-1}}=1,k=0,\pm1,\pm2,\cdots$$

所以若 n 阶复方阵 B 有 $\exp B=A$，则

$$\exp(B+2k\pi\sqrt{-1}I)=A,k=0,\pm1,\pm2,\cdots.$$

定理 13.3.4 设 A 为 n 阶非异复方阵，α 为非零复数，则存在 n 阶复方阵 B，使得 $B^\alpha=A$. 即矩阵方程 $X^\alpha=A$ 有解

证 今存在 n 阶复方阵 B_1，使得 $A=\exp B_1$. 令

$$B=\exp\left(\frac{1}{\alpha}B_1\right),$$

则有$B^\alpha=\left(\exp\left(\frac{1}{\alpha}B_1\right)\right)^\alpha=\exp B_1=A$. 证完.

注意：这里 B^α 理解为：无妨设 B 的特征根 λ_j 有$|\lambda_j-1|<1$,

$B^{\alpha}=\exp(\alpha\log B)$ 又上述矩阵方程的解不唯一，事实上，若 $B^{\alpha}=A$，则有

$$\left(B\exp\frac{2k\pi\sqrt{-1}}{\alpha}\right)^{\alpha}=B^{\alpha}\exp 2k\pi\sqrt{-1}=B^{\alpha}=A,$$

$$k=0,\pm 1,\pm 2,\cdots.$$

另外，对任意奇异方阵，是不可以任意开方的. 例如，不存在二阶复方阵$\begin{pmatrix}a & b\\ c & d\end{pmatrix}$，使得

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}^2=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}.$$

将上面理论推广，我们发现关键在于多项式函数和幂级数函数具有三条基本性质. 下面我们引进

定义 给定数值整函数 $f(z)$，即 $f(z)$在整个复平面上全纯. n^2个独立未知数 x_{ij} 构成之 n 阶方阵 $X=(x_{ij})$ 为“**自变量**”的函数 $f(X)$称为**纯函数**，如果它有性质：

(1) $f(P^{-1}XP)=P^{-1}f(X)P$ 对一切 n 阶非异常数复方阵 P 成立；

(2) $f(\mathrm{diag}(X_1^{(e_1)},\cdots,X_t^{(e_t)}))=\mathrm{diag}(f(X_1),\cdots,f(X_t))$;

(3) 取 $X=J_0^{(e_0)}=\lambda_0 I^{(e_0)}+N^{(e_0)}$为 Jordan 块，则有

$$f(J_0)=\sum_{l=0}^{e_0-1}\frac{1}{l!}f^{(l)}(\lambda_0)N^l$$

$$=\begin{pmatrix}f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \cdots & \frac{1}{(e_0-1)!}f^{(e_0-1)}(\lambda_0)\\ & f(\lambda_0) & \cdots & \frac{1}{(e_0-2)!}f^{(e_0-2)}(\lambda_0)\\ & & \ddots & \vdots\\ & & & f(\lambda_0)\end{pmatrix}.$$

即对角线上为 $f(\lambda_0)$，移上一排为$\frac{1}{1!}f'(\lambda_0)$，…，最后一排为

$\frac{1}{(e_0-1)!}f^{(e_0-1)}(\lambda_0)$.

引理 13.3.3 纯函数为单值函数，即纯函数 的定义有意义.

证 今任给 n 阶复方阵 A，则存在 n 阶非异复方阵 P，使得 $P^{-1}AP=J=\mathrm{diag}(J_1^{(e_1)},\cdots,J_t^{(e_t)})$ 为 Jordan 标准形，其中 $J_k^{(e_k)}$ 为由初等因式 $(\lambda-\lambda_k)^{e_k}$ 决定的 Jordan 块. 设 $f(X)$ 为纯函数，则在 A 的值为

$$f(A)=f(PJP^{-1})=Pf(J)P^{-1}=P\mathrm{diag}(f(J_1),\cdots,f(J_t))P^{-}$$

于是

$$f(A)=P\left[\mathrm{diag}\left(\sum_{l=0}^{e_1-1}\frac{1}{l!}f^{(l)}(\lambda_1)N^l,\cdots,\sum_{l=0}^{e_t-1}\frac{1}{l!}f^{(l)}(\lambda_t)N^l\right)\right]P^{-1}.$$

这样就算出了函数值. 但是 P 不唯一，即若另外存在一个 n 阶非异复方阵 Q，使得

$$A=PJP^{-1}=QJQ^{-1}.$$

我们要证明

$$Pf(J)P^{-1}=Qf(J)Q^{-1}.$$

于是在 $X=A$ 的函数值 $f(A)$ 与 P 的选取无关，即为单值函数.

事实上，由 $PJP^{-1}=QJQ^{-1}$，所以 $(Q^{-1}P)J=J(Q^{-1}P)$，将 n 阶非异复方阵 $R=Q^{-1}P$ 和 J 一样分块为

$$R=\begin{pmatrix}R_{11}&\cdots&R_{1t}\\\vdots&&\vdots\\R_{t1}&\cdots&R_{tt}\end{pmatrix},$$

于是 $$R_{ij}J_j=J_iR_{ij},\quad i,j=1,2,\cdots,t.$$

由定理 13.2.7，当 $\lambda_j\neq\lambda_i$，则有 $R_{ij}=0$；当 $\lambda_j=\lambda_i$，则

$$R_{ij}=\begin{pmatrix}0 & \sum_{l=0}^{\min(e_i,e_j)-1} b_l^{(ij)}N^l\\ 0 & 0\end{pmatrix}$$

其中

$$N=\begin{pmatrix}0 & 1 & & \\ & & \ddots & \\ & \ddots & & 1\\ & & & 0\end{pmatrix}$$

为 $\min(e_i, e_j)$阶方阵. 所以当 $\lambda_j \neq \lambda_i$ 时有

$$R_{ij}f(J_j)=f(J_i)R_{ij};$$

当 $\lambda_j=\lambda_i$ 时,

$$R_{ij}f(J_j)=R_{ij}\sum_{l=0}^{e_j-1}\frac{1}{l!}f^{(l)}(\lambda_j)N^l$$

$$=\sum_{l=0}^{e_j-1}\frac{1}{l!}f^{(l)}(\lambda_j)R_{ij}N^l.$$

当 $e_i \geqslant e_j$ 时

$$R_{ij}N^{(e_j)l}=\begin{pmatrix}\sum_{p=0}^{e_j-1}b_p^{(ij)}\ N^{(e_j)p}\\ 0\end{pmatrix}N^{(e_j)l}$$

$$=\begin{pmatrix}\sum_{p=0}^{e_j-1}b_p^{(ij)}\ N^{(e_j)(p+l)}\\ 0\end{pmatrix}=N^{(e_i)l}\begin{pmatrix}\sum_{p=0}^{e_j-1}b_p^{(ij)}\ N^{(e_j)p}\\ 0\end{pmatrix}$$

$$=N^{(e_i)l}R_{ij}.$$

同理, 当 $e_i<e_j$ 时

$$R_{ij}N^{(e_j)l}=\begin{pmatrix}0 & \sum_{p=0}^{e_i-1}b_p^{(ij)}\ N^{(e_i)p}\end{pmatrix}N^{(e_j)l}$$

$$=N^{(e_i)l}\begin{pmatrix}0 & \sum_{p=0}^{e_i-1}b_p^{(ij)}\ N^{(e_i)p}\end{pmatrix}=N^{(e_i)l}R_{ij}.$$

所以证明了当 $\lambda_j=\lambda_i$ 时, 有

$$R_{ij}f(J_j)=f(J_i)R_{ij}.$$

总之,

$$Rf(J)=f(J)R,$$

即有 $Q^{-1}Pf(J)=f(J)Q^{-1}P$，或 $Pf(J)P^{-1}=Qf(J)Q^{-1}$ 证完.

由定义立即有

定理13.3.5 设 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 n 阶复方阵 A 的 n 个特征根. 设 $f(X)$ 为纯函数，则 $f(A)$ 的 n 个特征根为 $f(\lambda_1)$，$f(\lambda_2)$，$\cdots$，$f(\lambda_n)$.

由于纯函数的定义依赖于所给的整函数 $f(z)$，所以 $\sin X$，$\cos X$，$\exp X$ 都是纯函数，而 $(I+X)^\alpha$，$\log(I+X)$，不是纯函数. 当然，我们也可以扩充纯函数的概念，使得"自变量"不能取任意方阵，而是限制在一定条件下的一批方阵. 换句话说，我们取一个具有定义域 $\mathscr{D}$ 的全纯函数 $f(z)$，只要方阵 A 的特征根必须落在定义域 $\mathscr{D}$ 中，那末矩阵函数 $f(X)$ 在 $X=A$，有意义. 在这样意义下的纯函数，就包括初等函数 $(I+X)^\alpha$，$\log(I+X)$ 了.

纯函数的基本性质依赖于下面引理.

引理13.3.4 （广义 Lagrange 插值公式）任给 sm 个复数 a_{jk}，$1\leqslant j\leqslant s$，$0\leqslant k<m$，任给 s 个不同复数 $\alpha_1,\alpha_2,\cdots,\alpha_s$，则存在多项式 $p(x)$，使得

$$p^{(k)}(\alpha_i)=a_{ik},\quad 1\leqslant i\leqslant s,\quad 0\leqslant k<m.$$

证 记

$$\Phi_j(\lambda)=(\lambda-\alpha_1)^m(\lambda-\alpha_2)^m\cdots(\lambda-\alpha_{j-1})^m(\lambda-\alpha_{j+1})^m\cdots(\lambda-\alpha_s)^m$$

令

$$\mu_{j0}=a_{j0}/\Phi_j(\alpha_j),\qquad j=1,2,\cdots,s,$$

作递推公式:

$$\mu_{jk}\Phi_j(\alpha_j)+\binom{k}{1}\mu_{j,k-1}\Phi_j'(\alpha_j)+\cdots+\binom{k}{k}\mu_{j0}\Phi_j^{(k)}(\alpha_j)=a_{jk},$$

$$k=0,1,2,\cdots$$

则决定了 $\mu_{j0},\mu_{j1},\mu_{j2},\cdots,\mu_{jm}$. 构造多项式

$$\varphi_j(x)=\mu_{j0}+\frac{1}{1!}\mu_{j1}(x-\alpha_j)+\cdots+\frac{1}{(m-1)!}u_{j,m-1}(x-\alpha_j)^{m-1}.$$

取 $j=1,2,\cdots,s$, 令

$$p(x)=\sum_{j=1}^{s}\Phi_j(x)\varphi_j(x).$$

我们来证明多项式 $p(x)$ 适合引理条件

取 $1\leqslant l\leqslant s, 0\leqslant k<m$, 则有

$$p^{(k)}(\alpha_l)=\frac{d^k}{d^kx}p(x)\bigg|_{x=\alpha_l}=\sum_{j=1}^{s}\left(\frac{d^k}{dx^k}(\Phi_j(x)\varphi_j(x)\right)_{x=\alpha_l}.$$

今取 $0\leqslant u\leqslant m$, 则

$$\begin{aligned}\frac{d^u}{dx^u}\Phi_j(x)\bigg|_{x=\alpha_l}&=\frac{d^u}{dx^u}\left(\prod_{v\neq j}(x-\alpha_v)^m\right|_{x=\alpha_l}\\&=\prod_{v\neq j}(x-\alpha_v)^{m-u}\cdot\xi_{ju}(x)\bigg|_{x=\alpha_l}\end{aligned}$$

其中 $\xi_{ju}(x)$ 为 x 的多项式. 所以当 $l\neq j$, 则有

$$\frac{d^u}{dx^u}\Phi_j(x)\bigg|_{x=\alpha_l}=0, 0\leqslant u\leqslant m.$$

所以

$$\begin{aligned}p^{(k)}(\alpha_l)&=\frac{d^k}{dx^k}(\Phi_l(x)\varphi_l(x))\bigg|_{x=\alpha_l}\\&=\sum_{v=0}^{k}\binom{k}{v}\Phi_l^{(v)}(\alpha_l)\varphi_l^{(k-v)}(\alpha_l)\\&=\sum_{v=0}^{k}\binom{k}{v}\mu_{l,k-v}\Phi_l^{(v)}(\alpha_l)=a_{lk}.\end{aligned}$$

证完.

定理 13.3.6 设 $f(X)$ 为方阵纯函数，且当 $X=A$ 时函数值 $f(A)$ 有意义, 则 $f(A)$ 可表为 A 的多项式.

证 符号同前

$A=PJP^{-1}=P\mathrm{diag}(J_1^{(e_1)},\cdots,J_t^{(e_t)})P_j^{-1},\ J_k=\lambda_k I^{(e_k)}+N^{(e_k)}.$

则

$$f(A)=Pf(J)P^{-1}=P\mathrm{diag}(f(J_1),\cdots,f(J_t))P^{-1}$$

今

$$f(J_k)=\sum_{l=0}^{e_k-1}\frac{1}{e_k!}f^{(l)}(\lambda_k)N^{(e_k l)}$$

在 $\lambda_1,\cdots,\lambda_t$ 中取出不同数 $\alpha_1,\cdots,\alpha_s$,. 令 $m=\max(e_1,\cdots,e_t)$. 由引理 13.3.4,存在多项式 $p(x)$,使得

$$p^{(l)}(\alpha_j)=f^{(l)}(\alpha_j),\ j=1,2\cdots,s,\qquad l=0,1,\cdots,m-1.$$

因此

$$\begin{aligned}f(J_k)&=\sum_{l=0}^{e_k-1}\frac{1}{l!}f^{(l)}(\lambda_k)N^{(e_k)l}\\&=\sum_{l=0}^{e_k-1}\frac{1}{l!}p^{(l)}(\lambda_k)N^{(e_k)l}\\&=p(J_k),\quad k=1,2,\cdots,t\end{aligned}$$

即

$$f(A)=P\ \mathrm{diag}(p(J_1),\cdots,p(J_t))P^{-1}=p(A).$$

定理证完.

这个定理有很多应用.

定理 13.3.7 对任一 n 阶非异复方阵 A,则存在 A 的多项式 $p(A)$,使得 $p(A)^2=A$.

证 今 $f(z)=\sqrt{z}$ 在 $z=0$ 点外面全纯. 而 A 非异, 所以特征根都不等于零,因此 $f(A)$ 有意义. 符号同上,

$$A=PJP^{-1}=P\mathrm{diag}(J_1^{(e_1)},\cdots,J_t^{(e_t)})P^{-1},$$
$$J_k=\lambda_k I^{(e_k)}+N^{(e_k)}.$$

于是

$$f(A)=P\operatorname{diag}\left(\sum_{l=0}^{e_1-1}\frac{1}{l!}f^{(l)}(\lambda_1)N^{(e_1)l},\cdots,\right.$$
$$\left.\sum_{l=0}^{e_t-1}\frac{1}{l!}f^{(l)}(\lambda_t)N^{(e_t)l}\right)P^{-1}.$$

而

$$f(A)^2=P\operatorname{diag}\left(\cdots,\left(\sum_{l=1}^{e_k-1}\frac{1}{l!}f^{(l)}(\lambda_k)N^{(e_k)l}\right)^2,\cdots\right)P^{-1},$$

又

$$\left(\sum_{l=0}^{e_k-1}\frac{1}{l!}f^{(l)}(\lambda_k)N^{(e_k)l}\right)^2=\sum_{l,j=0}^{\infty}\frac{1}{l!j!}f^{(l)}(\lambda_k)f^{(j)}(\lambda_k)N^{l+j}$$
$$=\sum_{u=0}^{\infty}\frac{1}{u!}\sum_{l+j=u}\binom{u}{l}f^{(l)}(\lambda_k)f^{(j)}(\lambda_k)N^{l+j}$$

$$=\sum_{u=0}^{\infty}\frac{1}{u!}\left(\frac{d^u}{dx^u}f(x)^2\Big|_{x=\lambda_k}\right)N^u.$$

由 $f(x)^2=x$，则有

$$f(A)^2=P\operatorname{diag}(\lambda_1I^{(e_1)}+N^{(e_1)},\cdots,\lambda_tI^{(e_t)}+N^{(e_t)})P^{-1}=A.$$

另一方面，由定理 13.3.6，存在多项式 $p(z)$，使得 $f(A)=p(A)$．这证明了 $p(A)^2=A$．定理证完．

定义 n 阶复方阵 A 和 B 称为**复正交相似的**，如果存在 n 阶**复正交方阵** O，即 $OO'=O'O=I^{(n)}$，使得

$$B=O^{-1}AO.$$

引理 13.3.5 复正交相似为等价关系．

定理 13.3.8 n 阶复对称(复斜对称、复正交)方阵如果相似于复对称(复斜对称、复正交)方阵，则它们必复正交相似．

证 由假设，存在 n 阶非异复方阵 P，使得

(a) $A'=A, (P^{-1}AP)'=P^{-1}AP$;

(b) $A'=-A, (P^{-1}AP)'=-P^{-1}AP$;

(c) $AA'=I, (P^{-1}AP)(P^{-1}AP)'=I$.

由直接计算可知这三种情形都导出

$$PP'A=APP'.$$

今 PP' 为非异复方阵，由定理 13.3.7，存在 PP' 的多项式 $p(PP')$ 使得 $p(PP')^2=PP'$. 自然 $p(PP')$ 也是非异复对称方阵，且

$$p(PP')A=Ap(PP').$$

记

$$O=P'(p(PP'))^{-1}$$

则有 $\quad P=(p(PP'))O'$

$$O'O=p(PP')^{-1}PP'p(PP')^{-1}=(PP')(p(PP'))^{-2}$$
$$=(PP')(PP')^{-1}=I,$$

即 O 为复正交方阵，而

$$B=P^{-1}AP=O(p(PP'))^{-1}A(p(PP'))O'=OAO'.$$

证完.

定理 13.3.9 任一 n 阶非异复方阵 A 必有分解

$$A=OS=S_1O,$$

其中 O 为复正交方阵，S 及 S_1 为 n 阶非异复对称方阵.

证 今 AA' 为 n 阶非异复对称方阵. 由定理 13.3.7，存在 AA' 的多项式 $p(AA')=S_1$，它仍为复对称方阵，且 $S_1^2=AA'$. 记 $O=S_1^{-1}A$，则

$$OO'=S_1^{-1}AA'S_1^{-1}=S_1^{-1}S_1^2S_1^{-1}=I,$$

所以 O 为复正交方阵，且有 $A=S_1O=O(O'S_1O)$，记 $S=O'S_1O$，它仍为复对称方阵. 所以有分解 $A=S_1O=OS$，证完.

定理 13.3.10 设 H 为 n 阶定正 Hermite 方阵，则存在复正交方阵 O，使得

$$O H \bar{O}' = \operatorname{diag}(\lambda_1, \cdots, \lambda_n),$$

其中 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n > 0$, 且 $\lambda_1^2, \cdots, \lambda_n^2$ 为 $H\bar{H}$ 的所有特征根

证 今 $H>0$, 所以存在 n 阶非异复方阵 Q, 使得

$$H = Q'\bar{Q}.$$

今 $QH\bar{Q}'>0$, 所以存在 n 阶酉方阵 U, 使得

$$UQH\bar{Q}'\bar{U}' = \operatorname{diag}(\mu_1 I^{(e_1)}, \cdots, \mu_t I^{(e_t)}) = \Lambda,$$

其中 $\mu_1 > \mu_2 > \cdots > \mu_t > 0$. 因此

$$H\bar{H} = [Q^{-1}\bar{U}'\Lambda U(\bar{Q}')^{-1}](\bar{Q}'Q) = Q^{-1}U^{-1}\Lambda UQ = (UQ)^{-1}\Lambda UQ.$$

所以 Λ 和 $H\bar{H}$ 相似, 今 $\Lambda' = \Lambda$, $(H\bar{H})' = \bar{H}'H' = H\bar{H}$, 即 Λ 和 $H\bar{H}$ 都是复对称方阵. 由定理 13.3.8, 存在复正交方阵 O_0 使得

$$H\bar{H} = O_0'\Lambda O_0.$$

于是

$$\begin{aligned} \Lambda &= O_0 H\bar{H} O_0' \\ &= (O_0 H \bar{O}_0')(O_0 H \bar{O}_0')' \end{aligned}$$

记 $H_0 = O_0 H \bar{O}_0'$, 则 H_0 为定正 Hermite 方阵, 且 $H_0 H_0' = \Lambda$, 所以 $H_0' = H_0^{-1}\Lambda$ 为 Hermite 方阵, 即有 $H_0^{-1}\Lambda = (\overline{H_0^{-1}\Lambda})' = \Lambda H_0^{-1}$, 即

$$H_0\Lambda = \Lambda H_0.$$

于是将 H_0 和 Λ 同样分块, 则有

$$H_0 = \operatorname{diag}(H_1^{(e_1)}, \cdots, H_t^{(e_t)}.$$

由 $\Lambda = H_0 H_0'$ 有

$$H_j H_j' = \mu_j I, \ j = 1, 2, \cdots, t.$$

于是

$$O_j^{(e_j)} = \frac{1}{\sqrt{\mu_j}} H_j$$

为 e_j 阶复正交方阵, 且为定正 Hermite 方阵, 将 O_j 之实部与虚部分开, 即记

$$O_j=S_j+\sqrt{-1}K_j,$$

则 S_j 实对称，K_j 实斜对称，且 $I=O_jO_j'=(S_j+\sqrt{-1}K_j)(S_j-\sqrt{-1}K_j)=S_j^2+K_j^2+\sqrt{-1}(K_jS_j-S_jK_j)$，即有

$$S_jK_j=K_jS_j,\ S_j^2+K_j^2=I.$$

而 $O_j>0$ 等价于实对称方阵

$$\begin{pmatrix} S_j & K_j \\ -K_j & S_j \end{pmatrix}>0.$$

于是不难证明存在实正交方阵 P_j，使得

$$O_j=P_j'\mathrm{diag}\left(\begin{pmatrix}\mathrm{ch}\theta_{1j} & \sqrt{-1}\mathrm{sh}\theta_{1j} \\ -\sqrt{-1}\mathrm{sh}\theta_{1j} & \mathrm{ch}\theta_{1j}\end{pmatrix}, \cdots, \begin{pmatrix}\mathrm{ch}\theta_{lj} & \sqrt{-1}\mathrm{sh}\theta_{lj} \\ -\sqrt{-1}\mathrm{sh}\theta_{lj} & \mathrm{ch}\theta_{lj}\end{pmatrix}, I\right)P_j,$$

取

$$Q_j=P_j'\mathrm{diag}\left(\begin{pmatrix}\mathrm{ch}\frac{1}{2}\theta_{1j} & \sqrt{-1}\mathrm{sh}\frac{1}{2}\theta_{1j} \\ -\sqrt{-1}\mathrm{sh}\frac{1}{2}\theta_{1j} & \mathrm{ch}\frac{1}{2}\theta_{1j}\end{pmatrix}, \cdots, \begin{pmatrix}\mathrm{ch}\frac{1}{2}\theta_{lj} & \sqrt{-1}\mathrm{sh}\frac{1}{2}\theta_{lj} \\ -\sqrt{-1}\mathrm{sh}\frac{1}{2}\theta_{lj} & \mathrm{ch}\frac{1}{2}\theta_{lj}\end{pmatrix}, I\right)P_j,$$

则 Q_j 仍为定正 Hermite 方阵，且 Q_j 复正交，又

$$Q_j^2=O_j,\ j=1,2,\cdots,t.$$

取

$$O=\mathrm{diag}(Q_1',\cdots,Q_t')O_0.$$

于是

$$\begin{aligned} OH\bar{O}' &= \mathrm{diag}(Q_1',\cdots,Q_t')O_0H\bar{O}_0'\mathrm{diag}(\bar{Q}_1,\cdots,\bar{Q}_t) \\ &= \mathrm{diag}(Q_1',\cdots,Q_t')H_0\mathrm{diag}(\bar{Q}_1,\cdots,\bar{Q}_t) \end{aligned}$$

$$=\operatorname{diag}(Q_1'H_1\bar{Q}_1,\cdots,Q_t'H_t\bar{Q}_t)$$
$$=\operatorname{diag}(\sqrt{\mu_1}Q_1'O_1\bar{Q}_1,\cdots,\sqrt{\mu_t}Q_t'O_t\bar{Q}_t).$$

今

$$Q_j'O_j\bar{Q}_j=P_j'P_j=I,j=1,2,\cdots,t.$$

因此

$$OH\bar{O}'=\operatorname{diag}(\sqrt{\mu_1}I^{(e_1)},\cdots,\sqrt{\mu_t}I^{(e_t)})=\Lambda^{\frac{1}{2}}.$$

这证明了定理．证完．

定理 13.3.11 和 n 阶复方阵 A 可交换的所有方阵可交换的方阵必为 A 的多项式．

证 考虑 n 阶 Jordan 标准形 $J=\operatorname{diag}(J_1^{(e_1)},\cdots,J_t^{(e_t)})$．由定理 13.2.7，和 J 可交换之方阵为

$$B=\begin{pmatrix}B_{11}&\cdots&B_{1t}\\ \vdots&&\vdots\\ B_{t1}&\cdots&B_{tt}\end{pmatrix},$$

其中 B_{ij} 为 $e_i\times e_j$ 矩阵．且当 $\lambda_i\neq\lambda_j$，则 $B_{ij}=0$；当 $\lambda_i=\lambda_j$，则

$$B_{ij}=\sum_{l=0}^{\min(e_i,e_j)-1}b_l^{(ij)}\begin{pmatrix}0&N^l\\0&0\end{pmatrix},$$

其中 N 为 $\min(e_i,e_j)$ 阶方阵 $\begin{pmatrix}0&1&&\\&\ddots&\ddots&\\&&&1\\&&&0\end{pmatrix}$，$b_l^{(ij)}$ 为任意常数．

设 n 阶方阵 C 有 $CB=BC$．将 C 和 B 一样分块为

$$C=\begin{pmatrix}C_{11}&\cdots&C_{1t}\\ \vdots&&\vdots\\ C_{t1}&\cdots&C_{tt}\end{pmatrix},$$

则有
$$\sum_{j=1}^{t}C_{ij}B_{jk}=\sum_{j=1}^{t}B_{ij}C_{jk},\quad i,k=1,2,\cdots,t.$$

由于 B_{ij}, B_{jk} 中元素分别由独立参数 $b_l^{(ij)}$，$b_l^{(jk)}$ 组成. 所以证明了

$$C_{jj}B_{jk}=B_{jk}C_{kk} \quad 1\leqslant j,k\leqslant t$$

$$C_{ij}B_{jk}=0, i\neq j; B_{ij}C_{jk}=0, j\neq k.$$

这证明了 $C_{ij}=0, i\neq j$, 而

$$C_{ii}=\sum_{l=0}^{e_i-1}\frac{1}{l!}a_{il}N^{(e_i)l}, i=1,2,\cdots,t.$$

无妨设 $e_i\leqslant e_j$, 设 $\lambda_i=\lambda_j$, 由 $B_{ij}C_{jj}=C_{ii}B_{ij}$ 可知

$$a_{il}=a_{jl}, l=0,1,\cdots,e_i-1.$$

由引理 13.3.4, 存在多项式 $p(x)$, 使得 $\left.\frac{d^l p(x)}{dx^l}\right|_{x=\lambda_i}=a_{il}$

所以

$$C_{ii}=\sum_{l=0}^{e_i-1}\frac{1}{l!}a_{il}N^{(e_i)l}=\sum_{l=0}^{e_i-1}\frac{1}{l!}\left.\frac{d^l p(x)}{dx^l}\right|_{x=\lambda_i}N^{(e_i)l}=p(J_i),$$

$i=1,2,\cdots,t$. 所以

$$C=\mathrm{diag}(c_{11},\cdots,c_{tt})=\mathrm{diag}(p(J_1),\cdots,p(J_t))=p(J).$$

今任取 n 阶方阵 A. 设 P 为 n 阶非异复方阵, 使得 $A=PJP^{-1}$, 其中 J 为 Jordan 标准形. 今

$$\mathfrak{S}=\{X \text{ 为 } n \text{ 阶复方阵}, XA=AX\},$$

所以

$$\mathfrak{S}=\{X \text{ 为 } n \text{ 阶复方阵}, XPJP^{-1}=PJP^{-1}X\}.$$

记

$$\mathfrak{S}_J=\{B \text{ 为 } n \text{ 阶复方阵}, BJ=JB\}$$

于是证明了

$$\mathfrak{S}_J=P^{-1}\mathfrak{S}P=\{P^{-1}XP\mid\forall X\in\mathfrak{S}\}.$$

今若 n 阶方阵 Y 和 $\mathfrak{S}$ 中任一方阵可交换, 则 $P^{-1}YP$ 和 $\mathfrak{S}_J$ 中任一方阵可交换, 所以 $P^{-1}YP=p(J)$, 其中 p 为多项式. 因此

$$Y=P(p(J))P^{-1}=p(PJP^{-1})=p(A).$$

这证明了定理.

定理 13.3.12(Jordan 分解) 设 A 为 n 阶复方阵,则有

$$A=S_1+N_1, S_1N_1=N_1S_1$$

其中 S_1 相似于对角形, N_1 为幂零方阵. 又 S_1 和 N_1 都是 A 的多项式,其中常数项都等于零. 又若

$$A=S_1+N_1=S_2+N_2, S_1N_1=N_1S_1, S_2N_2=N_2S_2,$$

其中 S_1 和 S_2 相似于对角形, N_1 和 N_2 为幂零方阵. 则有

$$S_2=S_1, N_2=N_1.$$

证 由于条件和结论都在相似下不变, 因此无妨设 A 为 Jordan 标准形 $\mathrm{diag}(J_1^{(e_1)}\cdots, J_t^{(e_t)})$,其中 $J_k^{(e_k)}$ 为由初等因式 $(\lambda-\lambda_k)^{e_k}$ 决定的 Jordan 块.

记 A 的不同特征根为 $\mu_1,\cdots,\mu_s$. 于是

$$\det(\lambda I-A)=\prod_{j=1}^{s}(\lambda-\mu_j)^{f_j}=\prod_{k=1}^{t}(\lambda-\lambda_k)^{e_k}.$$

当 $\mu_1,\cdots,\mu_s$ 中有一个为零时,取 $g_j(\lambda)=\det(\lambda I-A)(\lambda-\mu_j)^{-f_j}$;
当 $\mu_1,\cdots,\mu_s$ 都不等于零时,取 $g_j(\lambda)=\lambda\det(\lambda I-A)(\lambda-\mu_j)^{-f_j}$.
于是在 $\mu_j\neq0$ 时, $\lambda|g_j(\lambda)$. 另一方面,对 $l=1,2,\cdots,s$ 有 $(g_l(\lambda),(\lambda-\mu_l)^{f_l})=1$. 所以存在多项式 $u_l(\lambda),v_l(\lambda)$,使得

$$u_l(\lambda)g_l(\lambda)+v_l(\lambda)(\lambda-\mu_l)^{f_l}=1.$$

令

$$f(\lambda)=\sum_{l=1}^{s}\mu_l u_l(\lambda)g_l(\lambda).$$

由于当 $\mu_j=0$ 时, $\mu_j u_j(\lambda)g_j(\lambda)=0$; 当 $\mu_j\neq0$, $\lambda|g_j(\lambda)$. 所以 $\lambda|f(\lambda)$,即 $f(\lambda)$ 之常数项为零, 因此 $\lambda-f(\lambda)$ 的常数项也为零, 又有

$$(\lambda-\mu_j)^{f_j}|(f(\lambda)-\mu_j), j=1,2,\cdots,s.$$

令

$$S=f(A),\quad N=A-f(A)=A-S.$$

所以有

$$SN=NS, A=S+N.$$

因此为了证明这是 Jordan 分解，只需证明 S 相似于对角形，且 N 幂零．而

$$f(A)=f(\mathrm{diag}(J_1,\cdots,J_t))=\mathrm{diag}(f(J_1),\cdots,f(J_t)).$$

又由 $f(\lambda)=\mu_j+f_j(\lambda)(\lambda-\mu_j)^{f_j}$，其中 $f_j(\lambda)$ 为多项式．所以当 Jordan 块 J_k 有 $\lambda_k=\mu_j$，则 $e_k\leqslant f_j$．于是

$$\begin{aligned}f(J_k)&=\mu_jI+f_j(J_k)(J_k-\mu_jI)^{f_j}\\&=\lambda_kI+f_j(J_k)N^{(e_k)f_j}=\lambda_kI, k=1,2,\cdots,t.\end{aligned}$$

这证明了

$$\begin{aligned}S=f(A)&=\mathrm{diag}(\lambda_1I^{(e_1)},\cdots,\lambda_tI^{(e_t)}\\&=\mathrm{diag}(J_1^{(e_1)}-N^{(e_1)},\cdots,J_t^{(e_t)}-N^{(e_t)})\\&=A-\mathrm{diag}(N^{(e_1)},\cdots,N^{(e_t)}).\end{aligned}$$

所以 S 为对角形，而 $A-S=N=\mathrm{diag}(N^{(e_1)},\cdots,N^{(e_t)})$ 幂零．至此证明了分解的存在性．

下面证分解的唯一性．今若 $A=S_1+N_1=S_2+N_2$，其中 S_1，S_2 相似于对角形，N_1 和 N_2 幂零．于是 $S_1-S_2=N_2-N_1$．但是 $S_2N_2=N_2S_2$，所以 $S_2A=AS_2$．由于 S_1 为 A 的多项式，所以$S_1S_2=S_2S_1$．同理 $N_1N_2=N_2N_1$．于是 S_1-S_2 相似于对角形，且 N_2-N_1 幂零．这证明了 $S_1-S_2=0$，即 $S_2=S_1$．所以 $N_2=N_1$．唯一性证完．

证法二 记 $g(\lambda)=\det(\lambda I-A)$．作 $F(\lambda)=g(\lambda)(g,g')^{-1}$．则 $F(\lambda)=\prod\limits_{j=1}^{s}(\lambda-\mu_j)$，其中 $\mu_1,\cdots,\mu_s$ 为 A 的不同特征根．下面用归纳法来证明存在多项式 $h_k(\lambda)$，使得

$$F(\lambda)^{2^k} | F(\lambda - h_k(\lambda)F(\lambda)), k=1,2,\cdots.$$

事实上，由$(F, F')=1$，故存在多项式$u(\lambda), h_1(\lambda)$使得

$$u(\lambda)F(\lambda)+h_1(\lambda)F'(\lambda)=1.$$

而

$$F(\lambda-h_1(\lambda)F(\lambda))=F(\lambda)+F'(\lambda)(-h_1F)+*F^2$$
$$=F(\lambda)(1-h_1(\lambda)F'(\lambda))+*F^2$$

由 $F|(1-h_1F')$ 可知 $F^2|F(\lambda-h_1(\lambda)F(\lambda))$. 设存在多项式 $h_{k-1}(\lambda)$ 使得 $F(\lambda)2^{k-1}|F(\lambda-h_{k-1}(\lambda)F(\lambda))$. 记 $g_{k-1}(\lambda)=\lambda-h_{k-1}(\lambda)F(\lambda)$, 于是有 $F(\lambda)^{2^k}|F(g_{k-1})^2$, $u(g_{k-1})F(g_{k-1})+h_1(g_{k-1})F'(g_{k-1})=1$. 取

$$h_k(\lambda)=h_{k-1}(\lambda)+h_1(g_{k-1})F(g_{k-1})F(\lambda)^{-1}.$$

于是

$$F(\lambda-h_k(\lambda)F(\lambda))=F(g_{k-1}(\lambda)+(h_{k-1}(\lambda)-h_k(\lambda))F(\lambda))$$
$$=F(g_{k-1}(\lambda)-h_1(g_{k-1})F(g_{k-1}))$$
$$=F(g_{k-1})+F'(g_{k-1})(-h_1(g_{k-1})F(g_{k-1}))+F(g_{k-1})^2*$$
$$=F(g_{k-1})^2*.$$

这证明了 $F(\lambda)^{2^k}|F(\lambda-h_k(\lambda)F(\lambda))$.

今若 $F(0)\neq 0$，即方阵 A 的特征根不等于零，则可改取 $h_1(\lambda)$ 为 $h_1(\lambda)-F(0)^{-1}h_1(0)F(\lambda)$，于是有 $h_1(0)=0$. 因此若由归纳法假设， $h_{k-1}(0)=0$， 则 $h_k(0)=h_{k-1}(0)+h_1(-h_{k-1}(0)F(0))\cdot F(g_{k-1}(0))F(0)^{-1}=0$. 所以我们有 $F(0)=0$ 或者 $F(0)\neq 0$，则 $h_k(0)=0, k=1,2,\cdots$.

取最小自然数 m，使得 $F(x)|g(x), g(x)|F(x)^m$. 记自然数 k_0，有 $2^{k_0-1}<m\leqslant 2^{k_0}$. 记

$$N_1=h_{k_0}(A)F(A), S_1=A-N_1$$

自然 $S_1N_1=N_1S_1$，且 N_1 为 A 的多项式. 由于 $h_{k_0}(0)F(0)=0$,

所以 N_1 作为 A 的多项式,常数项为零. 同理, S_1 作为 A 的多项式,常数项为零. 为了证 $A=S_1+N_1$ 为 Jordan 分解, 只要证 S_1 相似于对角形,且 N_1 幂零就行了.

由 $g(x)|F(x)^m$,所以 $F(A)^m=0$,因此 $N_1^m=h_{k_0}(A)^mF(A)^m=0$,即 N_1幂零. 又 $S_1=A-N_1=A-h_{k_0}(A)F(A)$. 由$F(\lambda)^{2^{k_0}}|F(\lambda-h_{k_0}(\lambda)\ F(\lambda))$, 所以 $F(S_1)=F(A)^{2^{k_0}}\cdot h(A)$. 由 $m\leqslant 2^{k_0}$ 可知 $F(A)^{2^{k_0}}=0$. 因此证明了 $F(S_1)=0$, 即 $F(\lambda)$被 S_1 的极小多项式除尽. 但 $F(\lambda)$为 s 个一次因式之乘积,这证明了 S_1 的极小多项式无重根,所以 S_1 的初等因式都是一次的, 因此 S_1 相似于对角形. 证完.

注意 上面两种证明, 都是利用构造一些多项式, 从而造出 S_1 及 N_1. 因此,如果 A 为 n 阶实方阵, 则 S_1 及 N_1 仍为 n 阶实方阵. 所以 Jordan 分解在实的情形也成立.

证法三 下面的证明充分依赖于 Jordan 标准形, 无妨设 $A=\mathrm{diag}(J_1^{(e_1)},\ \cdots,\ J_t^{(e_t)})=\mathrm{diag}(\lambda_1I^{(e_1)}+N^{(e_1)},\ \cdots,\ \lambda_tI^{(e_t)}+N^{(e_t)})$为 Jordan 标准形. 取

$$S_1=\mathrm{diag}(\lambda_1I^{(e_1)},\cdots,\lambda_tI^{(e_t)}),\quad N_1=\mathrm{diag}(N^{(e_1)},\ \cdots,\ N^{(e_t)}).$$

自然 S_1 为对角形,N_1 为幂零方阵,又$S_1N_1=N_1S_1$. 余下证S_1 为 A 的多项式 $p(A)$,于是 $N_1=A-S_1=A-p(A)$仍为 A 的多项式.

今任取和 A 可交换的方阵 B,将B和 A 一样分块有

$$B=\begin{pmatrix}B_{11} & \cdots & B_{1t}\\ \vdots & & \vdots\\ B_{t1} & \cdots & B_{tt}\end{pmatrix}.$$

由定理 13.2.7,当 $\lambda_i\neq\lambda_j$,则 $B_{ij}=0$. 于是

$$S_1B-BS_1=((\lambda_i-\lambda_j)B_{ij})=0.$$

由定理 13.3.11,S_1 为 A 的多项式. 证完.

习题 13.3

1. 记 $X(t)=(x_{ij}(t))$，其中 $x_{ij}(t)$ 为 t 的可微函数.

$$\frac{d}{dt}X(t)=\left(\frac{d}{dt}x_{ij}(t)\right),$$

试求:

$$\frac{dX(t)}{dt}=X(t)A$$

的适合初值 $X(0)=I$ 的解，其中 A 和 $X(t)$ 之阶数相同.

2. 证明: A 为斜 Hermite 方阵当且仅当 $\exp(tA)$ 为酉方阵; 又证: A 为复斜对称方阵当且仅当 $\exp(tA)$ 为复正交方阵, $-\infty<t<\infty$.

3. 试证:

$$\exp\begin{pmatrix} 0 & \lambda \\ -\lambda & 0 \end{pmatrix}=\begin{pmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{pmatrix}.$$

4. 试证: 对任一 n 阶复方阵 A，总有

$$\cos^2 A+\sin^2 A=I.$$

5. 设 A 为特征根之模小于 1 的 n 阶复方阵. 试证:

$$\exp(\log(I+A))=I+A.$$

6. 设 O 为 n 阶正交定正 Hermite 方阵，试证: 存在实斜对称方阵 K，使得 $O=\exp(\sqrt{-1}K)$.

7. 设 $f_i(x)$ 为光滑函数，$i=1,2$，X 为 n^2 个独立未知数构成的 n 阶方阵. 记 $f_i(X)$ 为纯函数，$i=1, 2$. 试证: $f_i(X')=f_i(X)'$; 又记 $f_3(x)=f_1(x)+f_2(x)$，$f_4(x)=f_1(x)f_2(x)$，则纯函数 $f_3(X)$，$f_4(X)$ 仍有意义，且 $f_3(X)=f_1(X)+f_2(X)$，$f_4(X)=f_1(X)f_2(X)$.

8. 试证: 由 $f(x)=x^{-1}$ 定义的纯函数 $f(X)$ 只有在 n 阶方阵 $X=A$ 非异时才有意义，这时 $f(A)=A^{-1}$. 又由 $f(x)=x^{a/b}$ 定义的纯函数 $f(X)$，它在有意义的 n 阶方阵 $X=A$ 上，总有 $f(A)^b=A^a$，其中 a,b 为自然数。

9. 试证: n 阶非异复方阵 A 的逆方阵 A^{-1} 为 A 的多项式. 又对任意非零整数 l，则存在 A 的多项式 $p_l(A)$，使得 $p_l(A)^l=A$. 特别，当 A 复对称时，$S_l=p_l(A)$ 也复对称.

10. 试证: n 阶复方阵

$$\lambda_0 I+S,\quad S=\begin{pmatrix}0 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 0\end{pmatrix}+\sqrt{-1}\begin{pmatrix} & & 1 & 0 \\ & ⋰ & ⋰ & -1 \\ 1 & ⋰ & ⋰ & \\ 0 & -1 & & \end{pmatrix}$$

的初等因式组为$(\lambda-\lambda_0)^n$,所以相似于 Jordan 块 $\lambda_0 I^{(n)}+N^{(n)}$. 因此有

(1) 任一 n 阶复方阵相似于复对称方阵;

(2) 任一 n 阶复对称方阵复正交相似于

$$\operatorname{diag}(\lambda_1 I+S^{(e_1)},\cdots,\lambda_t I+S^{(e_t)}),$$

其中 S 定义如上.

§13.4 复方阵在复相似下的标准形

定义 n 阶复方阵 A 和 B 称为**复相似的**,如果存在 n 阶非异复方阵 P,使得

$$B=PA\bar{P}^{-1}.$$

引理 13.4.1 复相似为等阶关系.

引理 13.4.2 设 n 阶复方阵 A 和 B 复相似,则 $A\bar{A}$ 和 $B\bar{B}$ 相似.

证 今 $B=PA\bar{P}^{-1}$,其中 P 为 n 阶非异复方阵. 所以

$$B\bar{B}=(PA\bar{P}^{-1})(\bar{P}\bar{A}P^{-1})=PA\bar{A}P^{-1}.$$

证完.

这引理的逆不成立. 例如,取

$$A=\begin{pmatrix}\sqrt{-1} & 1 \\ -1 & \sqrt{-1}\end{pmatrix},\quad B=0,$$

则 $A\bar{A}=B\bar{B}=0$, 但是 B 复相似于零, 所以 A 和 B 不可能复相似.

由引理 13.4.2,在给定 n 阶复方阵 A 后, 我们先考虑 n 阶复方阵$A\bar{A}$,所以存在 n 阶复方阵 P 使得

$$PA\bar{A}P^{-1}=\operatorname{diag}(J_1^{(e_1)}\cdots,\ J_t^{(e_t)})=J,\quad J_k^{(e_k)}=\lambda_k I^{(e_k)}+N^{(e_k)}$$

为 Jordan 标准形. 记

$$B=PA\bar{P}^{-1},$$

于是有 $B\bar{B}=\mathrm{diag}(J_1,\cdots,J_t)=J$. 将 B 和 J 一样分块为

$$B=\begin{pmatrix} B_{11}\cdots B_{1t} \\ \vdots \qquad \vdots \\ B_{t1}\cdots B_{tt}\end{pmatrix}.$$

则由

$$(B\bar{B})B=B(\bar{B}B)=B(\overline{\bar{B}\bar{B}})$$

所以有 $JB=B\bar{J}$, 即有

$$J_jB_{jk}=B_{jk}\bar{J}_k,\quad j,k=1,2,\cdots,t.$$

由定理 13.2.7, 所以当 $\lambda_j\neq\bar{\lambda}_k$, 则有 $B_{jk}=0$; 当 $\lambda_j=\bar{\lambda}_k$, 则有

$$B_{jk}^{(e_j,e_k)}=\begin{pmatrix} 0 & \sum\limits_{l=1}^{\min(e_j,e_k)-1} b_{jk}^{(l)}N^{(\min(e_j,e_k))l} \\ 0 & 0\end{pmatrix}.$$

我们将 Jordan 标准形 $J=\mathrm{diag}(J_1,\cdots,J_t)$ 的 Jordan 块按如下方式排列: 特征根为零的 Jordan 块放在一起依次排下去, 记作 A_0; 特征根为非零实数的 Jordan 块放在一起依次排下去, 分别记作 $A_1,\cdots,A_p$. 其中 A_j 中 Jordan 块的特征根相同; 特征根为复数, 但不是实数的 Jordan 块放在一起, 接下去排它的共轭复数为特征根的 Jordan 块, 记作 $A_{p+1},\cdots,A_s$, 其中 A_k 中 Jordan 块的特征根为复数及其共轭复数. 于是

$$J=\mathrm{diag}(A_0,A_1,\cdots,A_p,A_{p+1},\cdots,A_s)$$

且相应方阵

$$B=\mathrm{diag}(B_0,B_1,\cdots,B_p,B_{p+1},\cdots,B_s).$$

有

$$B_j\bar{B}_j=A_j, j=0,1,\cdots,s.$$

因此为了求复方阵在复相似下标准形, 无妨设 Jordan 标准形 $J=$

$\mathrm{diag}(J_1,\cdots,J_t)$的特征根只有 λ_0 及 $\bar{\lambda}_0$. 下面分别在 $\lambda_0=0$ 及 $\lambda_0\neq 0$ 两种情形讨论. 在前一情形, B 奇异, 在后一情形, B 非异.

(一) $\lambda_0=0$, 于是 $J=\mathrm{diag}(N^{(e_1)},\cdots,N^{(e_t)})$, $e_1\geqslant\cdots\geqslant e_t>0$. 这时 $B\bar{B}=J$ 有初等因式组 $\lambda^{e_1},\lambda^{e_2},\cdots,\lambda^{e_t}$.

定理 13.4.1 条件如上, 则 B 复相似于唯一的标准形

$$\mathrm{diag}(N^{(f_1)},\cdots,N^{(f_s)}),\ f_1\geqslant\cdots\geqslant f_s>0.$$

证 对阶数 n 作归纳法. 当 $n=1$, 则 $B\bar{B}=J=0$, 于是 $B=0$. 设对 $n-1$ 阶复方阵 B_1, 若适合条件, 则定理成立. 今考虑 n 阶复方阵 B. 由 $\det B=0$, 所以 B 有属于特征根零的特征向量 α_1. 因此可构造 n 阶非异复方阵 $P_1=(\bar{\alpha}_1\bar{\alpha}_2\cdots\bar{\alpha}_n)^{-1}$, 这时

$$B_0=P_1B\bar{P}_1^{-1}=P_1B(\alpha_1\alpha_2\cdots\alpha_n)=P_1(0\quad *)$$

$$=\begin{pmatrix}0 & \beta' \\ 0 & B_1^{(n-1)}\end{pmatrix}.$$

而

$$B_0\bar{B}_0=\begin{pmatrix}0 & \beta'\bar{B}_1 \\ 0 & B_1\bar{B}_1\end{pmatrix}=P_1B\bar{B}P_1^{-1}=P_1\mathrm{diag}(N^{(e_1)},\cdots,N^{(e_t)})P_1^{-1}.$$

于是 $B_1\bar{B}_1$ 仍为幂零方阵. 由归纳法假设, 存在 $n-1$ 阶非异复方阵 P_2 使得

$$P_2B_1\bar{P}_2^{-1}=\mathrm{diag}(N^{(q_1)},\cdots,N^{(q_r)}).$$

于是

$$\begin{pmatrix}1 & 0 \\ 0 & P_2\end{pmatrix}B_0\overline{\begin{pmatrix}1 & 0 \\ 0 & P_2\end{pmatrix}}^{-1}=\begin{pmatrix}0 & \beta_1' & \cdots & \beta_r' \\ 0 & N^{(q_1)} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & N^{(q_r)}\end{pmatrix},$$

其中 $(\beta_1'\cdots\beta_r')=\beta'\bar{P}_2^{-1}$. 再作复相似

$$\begin{pmatrix}1 & (-\beta_1'N^{(q_1)'} & \cdots & -\beta_r'N^{(q_r)'}) \\ 0 & & I & \end{pmatrix}\begin{pmatrix}0 & (\beta_1' \ \cdots \ \beta_r') \\ 0 & \mathrm{diag}(N^{(q_1)},\cdots,N^{(q_r)})\end{pmatrix}$$

$$\cdot\begin{pmatrix}1 & (-\beta_1' N^{(q_1)'} & \cdots & -\beta_r' N^{(q_r)'}) \\ 0 & & I & \end{pmatrix}^{-1}$$

$$=\begin{pmatrix}0 & \beta_1'(I-N^{(q_1)'}N^{(q_1)}),\cdots & \beta_r'(I-N^{(q_r)'}N^{(q_r)}) \\ 0 & \mathrm{diag}(N^{(q_1)},\cdots,N^{(q_r)}) & \end{pmatrix}$$

由于 $N^{(q)'}N^{(q)}=\begin{pmatrix}0 & 0 \\ 0 & I^{(q-1)}\end{pmatrix}$，所以记

$$\delta_j'=\beta_j'(I-N(^{(q_j)'}N^{(q_j)})=\beta_j'\begin{pmatrix}1 & 0 \\ 0 & 0^{(q_j-1)}\end{pmatrix}=a_j(1,0,\cdots,0),$$

$$j=1,2,\cdots,r.$$

这证明了 B 复相似于

$$B_1=\begin{pmatrix}0 & a_1e_1^{(q_1)} & \cdots & a_re_1^{(q_r)} \\ 0 & N^{(q_1)} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & N^{(q_r)}\end{pmatrix},$$

其中 $e_i^{(q)}=(1,0,\cdots,0)\in R^q$. 如果 $(a_1,\cdots,a_r)=0$，则 B 复相似于 $\mathrm{diag}(N^{(1)},N^{(q_1)},\ \cdots,\ N^{(q_r)})$，所以证明了标准形存在性. 如果 $(a_1,\cdots,a_r)\neq 0$. 对 B_1 的行及列作同步的改变，从而无妨设 $a_1\neq 0$，$\cdots,a_v\neq 0,a_{v+1}=\cdots=a_r=0$. 记

$$P_3=\begin{pmatrix}1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \Lambda_1 & \Lambda_2\Lambda_2 & \cdots & X_v\Lambda_v & 0 \\ 0 & 0 & \Lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \Lambda_v & 0 \\ 0 & 0 & 0 & \cdots & 0 & I\end{pmatrix},$$

其中 $\Lambda_j^{(q_j)}=\mathrm{diag}(\bar{a}_j,a_j,\bar{a}_j,\cdots)$，于是 $\Lambda_j^{(q_j)}N^{(q_j)}=N^{(q_j)}\overline{\Lambda_j^{(q_j)}}$，$j=1$，$2,\cdots,v$. 因此

$$P_3\begin{pmatrix}0 & (a_1e_1^{(q_1)}\cdots a_1e_1^{(q_r)}) \\ 0 & \mathrm{diag}(N^{(q_1)},\cdots,N^{(q_r)})\end{pmatrix}=\bar{P}_3^{-1}$$

$$\begin{pmatrix} 0 & e_1^{(q_1)'} & e_1^{(q_2)'}-e_1^{(q_1)'}\overline{X}_2 & \cdots & e_1^{(q_v)'}-e_1^{(q_1)'}\overline{X}_v & \\ 0 & N^{(q_1)} & X_2N^{(q_2)}-N^{(q_1)}\overline{X}_2 & \cdots & X_vN^{(q_v)}-N^{(q_1)}\overline{X}_v & 0 \\ 0 & 0 & N^{(q_2)} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & N^{(q_v)} & 0 \\ 0 & 0 & 0 & \cdots & 0 & \mathrm{diag}(N^{(q_v+1)},\cdots,N^{(q_r)}) \end{pmatrix}$$

取 $X_j=\begin{pmatrix} I \\ 0 \end{pmatrix}$或$(I0)$，于是 $\overline{X}_j=\overline{X}_j, \overline{X}_jN^{(q_j)}=N^{(q_j)}\overline{X}_j,\ 1\leqslant j\leqslant v.$

这证明了

$$\begin{aligned} &P_3\begin{pmatrix} 0 & (a_1e_1^{(q_1)'}\cdots a_re_1^{(q_r)'}) \\ 0 & \mathrm{diag}(N^{(q_1)},\cdots,N^{(q_r)}) \end{pmatrix}\overline{P}_3^{-1} \\ &=\begin{pmatrix} 0 & (1\ 0\ \cdots\ 0) \\ 0 & \mathrm{diag}(N^{(q_1)},\cdots,N^{(q_r)}) \end{pmatrix} \\ &=\mathrm{diag}(N^{(q_1+1)},N^{(q_2)},\cdots,N^{(q_r)}). \end{aligned}$$

所以证明了标准形的存在性.

下面证明标准形的唯一性. 设 $f_1\geqslant\cdots\geqslant f_s>0$, $g_1\geqslant\cdots\geqslant g_r>0$，且 n 阶方阵

$$D_1=\mathrm{diag}(N^{(f_1)},\cdots,N^{(f_s)}), D_2=\mathrm{diag}(N^{(g_1)},\cdots,N^{(g_r)})$$

复相似. 所以存在非异复方阵 $P=X+\sqrt{-1}Y$，使得 $D_2=PD_1\overline{P}^{-1}$，其中 X,Y 为 n 阶实方阵. 因此有

$$D_2X=XD_1, -D_2Y=YD_1,$$

即

$$\begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}\begin{pmatrix} D_1 & 0 \\ 0 & -D_1 \end{pmatrix}=\begin{pmatrix} D_2 & 0 \\ 0 & -D_2 \end{pmatrix}\begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}.$$

今

$$\det\begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}=\det\left[\frac{1}{\sqrt{2}}\begin{pmatrix} I & -\sqrt{-1}I \\ I & \sqrt{-1}I \end{pmatrix}\right]\begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}$$

$$\cdot\left[\frac{1}{\sqrt{2}}\begin{pmatrix} I & -\sqrt{-1}I \\ I & \sqrt{-1}I \end{pmatrix}\right]'$$

$$=\det\begin{pmatrix} P & 0 \\ 0 & \bar{P} \end{pmatrix}=|\det P|^2>0.$$

这证明了 $\mathrm{diag}(D_1,-D_1)$和 $\mathrm{diag}(D_2,-D_2)$实相似，因此它们的初等因式组$\{\lambda^{f_1},\cdots,\lambda^{f_s},\lambda^{f_1},\cdots,\lambda^{f_s}\}$和$\{\lambda^{g_1},\cdots,\lambda^{g_r},\ \lambda^{g_1},\cdots,\lambda^{g_r}\}$为同一集合，所以 $s=r, f_i=g_i, i=1,2,\cdots,r$. 这也证明了 $D_2=D_1$. 所以标准形在复相似下唯一. 证完.

(二) $\lambda_0\neq 0$，于是 J 非异，所以 B 非异.

定理 13.4.2 设 A 和 B 为 n 阶非异复方阵，则 A 和 B 复相似当且仅当 $A\bar{A}$ 和 $B\bar{B}$ 相似.

证 由引理 14.4.2 可知只要证明：若存在 n 阶非异复方阵 P，使得 $PA\bar{A}=B\bar{B}P$，则 A 和 B 复相似. 事实上，

$$(B\bar{B}P)A=B(\bar{B}PA)=B(\overline{B\bar{P}\bar{A}}),$$

$$(B\bar{P}\bar{A})A=B(\bar{P}\bar{A}A)=B(\overline{B\bar{B}P}).$$

引进实参数 θ，记

$$Q(\theta)=(B\bar{B}P)e^{\sqrt{-1}\theta}+(B\bar{P}\bar{A})e^{-\sqrt{-1}\theta}.$$

则有

$$Q(\theta)A=B\overline{Q(\theta)}.$$

而

$$\det Q(\theta)=\det(B\bar{B}Pe^{\sqrt{-1}\theta}+B\bar{P}\bar{A}e^{-\sqrt{-1}\theta})$$

$$=\det\ e^{-\sqrt{-1}\theta}B\bar{B}P\ \det(e^{2\sqrt{-1}\theta}I^{(n)}+P^{-1}\bar{B}^{-1}\bar{P}\bar{A}).$$

由此可知，存在实数 θ，使得 $\det Q(\theta)\neq 0$. 这证明了 A 和 B 复相似. 证完.

由此可知，当 $\lambda_0\neq 0$ 时，我们只要具体构造一个 n 阶复方阵 C，它具有最简单的形式，且 $C\bar{C}$ 和 $A\bar{A}$ 相似，即有相同初等因式组即可.

(a) 设 $\bar{\lambda}_0 \neq \lambda_0$. 于是在复相似意义下可取

$B\bar{B} = \mathrm{diag}(\lambda_0 M^{(e_1)}, \cdots, \lambda_0 M^{(e_t)}, \bar{\lambda}_0 M^{(f_1)}, \cdots, \bar{\lambda}_0 M^{(f_s)})$, 其中 $M = I + N$. 所以 $B\bar{B}$ 的初等因式组为

$$(\lambda-\lambda_0)^{e_1}, \cdots, (\lambda-\lambda_0)^{e_t}, (\lambda-\bar{\lambda}_0)^{f_1}, \cdots, (\lambda-\bar{\lambda}_0)^{f_s}.$$

定理 13.4.3 条件同上，这时 B 复相似于标准形

$$C = \mathrm{diag}\left(\begin{pmatrix} 0 & \sqrt{\bar{\lambda}_0} M^{(e_1)} \\ \sqrt{\bar{\lambda}_0} M^{(e_1)} & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & \sqrt{\bar{\lambda}_0} M^{(e_t)} \\ \sqrt{\bar{\lambda}_0} M^{(e_t)} & 0 \end{pmatrix}\right),$$

所以初等因式组为

$$(\lambda-\lambda_0)^{e_1}, (\lambda-\bar{\lambda}_0)^{e_1}, \cdots, (\lambda-\lambda_0)^{e_t}, (\lambda-\bar{\lambda}_0)^{e_t}.$$

证 今 $B\bar{B} = B(\bar{B}B)B^{-1}$, 所以 $B\bar{B}$ 和 $\overline{B\bar{B}}$ 相似. 已知 $B\bar{B}$ 之初等因式组为$(\lambda-\lambda_0)^{e_1}, \cdots, (\lambda-\lambda_0)^{e_t}$, $(\lambda-\bar{\lambda}_0)^{f_1}, \cdots, (\lambda-\bar{\lambda}_0)^{f_s}$, 于是 $\overline{B\bar{B}}$ 之初等因式组为$(\lambda-\bar{\lambda}_0)^{e_1}$, $\cdots$, $(\lambda-\bar{\lambda}_0)^{e_t}$, $(\lambda-\lambda_0)^{f_1}$, $\cdots$, $(\lambda-\lambda_0)^{f_s}$. 由于 $\lambda_0 \neq \bar{\lambda}_0$, 这证明了 $s=t$, $f_i = e_i$, $i=1,2,\cdots,t$. 另一方面

$$C\bar{C} = \mathrm{diag}(\lambda_0 M^{(e_1)2}, \bar{\lambda}_0 M^{(e_1)2}, \cdots, \lambda_0 M^{(e_t)2}, \bar{\lambda}_0 M^{(e_t)2}).$$

而 $M^{(e)2} = (I^{(e)} + N^{(e)})^2 = I^{(e)} + 2N^{(e)} + N^{(e)2}$, 它的初等因式组为 $(\lambda-1)^e$. 这证明了 $C\bar{C}$ 之初等因式组为$(\lambda-\lambda_0)^{e_1}, (\lambda-\bar{\lambda}_0)^{e_1}, \cdots, (\lambda-\lambda_0)^{e_t}, (\lambda-\bar{\lambda}_0)^{e_t}$, 所以 $C\bar{C}$ 和 $B\bar{B}$ 有相同的初等因式组. 由定理 13.4.2, B 和 C 复相似. 证完.

(b) 设 $\bar{\lambda}_0 = \lambda_0 \neq 0$, 于是可取 Jordan 标准形

$$B\bar{B} = \lambda_0 \mathrm{diag}(M^{(e_1)}, \cdots, M^{(e_t)}), e_1 \geqslant \cdots \geqslant e_t > 0.$$

所以 $B\bar{B}$ 有初等因式组

$$(\lambda-\lambda_0)^{e_1}, \cdots, (\lambda-\lambda_0)^{e_t}.$$

定理 13.4.4 设 $\bar{\lambda}_0 = \lambda_0 > 0$, 则 B 复相似于标准形

$$C = \sqrt{\lambda_0} \mathrm{diag}(M^{(e_1)}, \cdots, M^{(e_t)}), e_1 \geqslant \cdots \geqslant e_t > 0.$$

证 今

$$C\bar{C}=\lambda_0\mathrm{diag}(M^{(e_1)2},\cdots,M^{(e_t)2}),$$

它的初等因式组为$(\lambda-\lambda_0)^{e_1},\cdots,(\lambda-\lambda_0)^{e_t}$，所以 $C\bar{C}$ 和 $B\bar{B}$ 相似. 由定理 13.4.2，所以 B 和 C 复相似. 证完.

定理 13.4.5 设 $\bar{\lambda}_0=\lambda_0<0$，则 B 复相似于标准形

$$C=\sqrt{-\lambda_0}\mathrm{diag}\left(\begin{pmatrix}0 & M^{(f_1)}\\ -M^{(f_1)} & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & M^{(f_s)}\\ -M^{(f_s)} & 0\end{pmatrix}\right),$$
$$f_1\geqslant\cdots\geqslant f_s>0,$$

其中

$$t=2s, f_1=e_1=e_2,\cdots,f_s=e_{2s-1}=e_{2s}.$$

证 令

$$C\bar{C}=\lambda_0\mathrm{diag}(M^{(f_1)2},M^{(f_1)2},\cdots,M^{(f_s)2},M^{(f_s)2}).$$

它的初等因式组为$(\lambda-\lambda_0)^{f_1}$，$(\lambda-\lambda_0)^{f_1},\cdots,(\lambda-\lambda_0)^{f_s}$，$(\lambda-\lambda_0)^{f_s}$. 由定理 13.4.2，为了证明定理，只要证 $t=2s, e_1=e_2,\cdots,e_{2s-1}=e_{2s}$ 就够了. 即设

$$e_1\geqslant\cdots\geqslant e_l>e_{l+1}=\cdots=e_{l+p}>e_{l+p+1}\geqslant\cdots\geqslant e_t>0,$$

则要证 p 为偶数.

事实上，由 $B\bar{B}=\lambda_0\mathrm{diag}(M^{(e_1)},\cdots,M^{(e_t)})$. 将 B 和 $B\bar{B}$ 一样分块为

$$B=\begin{pmatrix}B_{11} & \cdots & B_{1t}\\ \vdots & & \vdots\\ B_{t1} & \cdots & B_{tt}\end{pmatrix},$$

于是由 $B(\bar{B}B)=(B\bar{B})B$，即 $B(B\bar{B})=(B\bar{B})B$，有

$$M^{(e_i)}B_{ij}=B_{ij}M^{(e_j)},\quad i,j=1,2,\cdots,t.$$

但是 $M^{(e_i)}=I^{(e_i)}+N^{(e_i)}$，$M^{(e_j)}=I^{(e_j)}+N^{(e_j)}$，所以有

$$N^{(e_i)}B_{ij}=B_{ij}N^{(e_j)},\quad i,j=1,2,\cdots,t.$$

由定理 13.2.7 有

$$B_{ij}^{(e_i,e_j)}=\sum_{q=0}^{\min(e_i,e_j)-1} b_{ij}^{(q)}\begin{pmatrix}0 & N^q\\ 0 & 0\end{pmatrix},$$

其中$N=N^{(\min(e_i,e_j))}$为$\min(e_i,e_j)$阶方阵$\begin{pmatrix}0 & 1 & & \\ & \ddots & \ddots & \\ & & \theta & 1\end{pmatrix}$. 但是由 $B\bar{B}=\lambda_0 \mathrm{diag}(M^{(e_1)},\cdots,M^{(e_l)})$, 所以有

$$\sum_{j=1}^{l} B_{ij}\bar{B}_{jk}=\delta_{ik}\lambda_0 M^{(e_i)}=\delta_{ik}\lambda_0(I^{(e_i)}+N^{(e_i)}).$$

考虑 $i,k=l+1,l+2,\cdots,l+p$. 当 $j\neq l+1,l+2,\cdots,l+p$, 则有 $e_j\neq e_i=e_k$. 这时 $B_{ij}\bar{B}_{jk}$ 的对角元素为零, 所以只要考虑

$$\sum_{j+l=1}^{l+p} B_{ij}\bar{B}_{jk}=\sum_{r,s=0}^{e_l-1}\sum_{j=l+1}^{l+p} b_{ij}^{(r)}\overline{b_{jk}^{(s)}}N^{r+s},$$

它的对角元素为$\sum_{i=l+1}^{l+p} b_{ij}^{(0)}\overline{b_{jk}^{(0)}}$. 这证明了

$$\sum_{j=l+1}^{l+p} b_{ij}^{(0)}b_{jk}^{(0)}=\delta_{ik}\lambda_0,\quad i,k=l+1,\cdots,l+p.$$

所以 p 阶方阵.

$$Q=\begin{pmatrix} b_{l+1,l+1}^{(0)} & \cdots & b_{l+1,l+p}^{(0)}\\ \vdots & & \vdots\\ b_{l+p,l+1}^{(0)} & \cdots & b_{l+p,l+p}^{(0)}\end{pmatrix}$$

有 $Q\bar{Q}=\lambda_0 I^{(p)}$双方取行列式, 有 $\lambda_0^p=|\det Q|^2>0$. 但是$\lambda_0<0$, 这证明了 p 为偶数. 证完.

最后给出复相似的一个应用.

定理 13.4.6 设 A 为 n 阶复方阵, 则存在 Hermite 方阵 H_1 和 H_2, 及复对称方阵 S_1 和 S_2, 使得

$$A=H_1S_1=S_2H_2.$$

且可事先规定哪一个是非异的.

证 显然只证 $A=H_1S_1$ 即可. 今若有分解 $A=H_1S_1$, 则对任

一 n 阶复非异方阵 P 有

$$PA\overline{P}^{-1}=PH_1\overline{P}'(\overline{P^{-1}})'S_1(\overline{P^{-1}}).$$

而 $PH_1\overline{P}'$ 仍为 Hermite 方阵，$(\overline{P^{-1}})'S_1(\overline{P^{-1}})$ 仍为复对称方阵. 所以问题化为对复方阵在复相似下的标准形来证明定理.

记

$$N=\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{pmatrix},\ M=I+N,\ \widetilde{M}=\begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}.$$

则

$$N\widetilde{M}=\begin{pmatrix} & & 1 & 0 \\ & \iddots & \iddots & \\ 1 & \iddots & & \\ 0 & & & \end{pmatrix}.$$

当 $\lambda_0=0$，作 $N=(N\widetilde{M})\widetilde{M}$；当 $\lambda_0>0$，作 $\sqrt{\lambda_0}M=\sqrt{\lambda_0}(M\widetilde{M})\widetilde{M}$；当 $\lambda_0<0$，作 $\sqrt{-\lambda_0}\begin{pmatrix} 0 & M \\ -M & 0 \end{pmatrix}=\sqrt{-\lambda_0}\begin{pmatrix} 0 & M\widetilde{M} \\ -M\widetilde{M} & 0 \end{pmatrix}\begin{pmatrix} \widetilde{M} & 0 \\ 0 & \widetilde{M} \end{pmatrix}$；

当 $\lambda_0\neq\overline{\lambda}_0$，作

$$\begin{pmatrix} 0 & \sqrt{\overline{\lambda}_0}M \\ \sqrt{\overline{\lambda}_0}M & 0 \end{pmatrix}=\begin{pmatrix} 0 & \sqrt{\overline{\lambda}_0}M\widetilde{M} \\ \sqrt{\overline{\lambda}_0}M\widetilde{M} & 0 \end{pmatrix}\begin{pmatrix} \widetilde{M} & 0 \\ 0 & \widetilde{M} \end{pmatrix}.$$

由直接验证，便证明了定理.

习题 13.4

1. 试证：两个 n 阶复方阵 A 和 B 复相似当且仅当 $2n$ 阶复方阵 $\begin{pmatrix} 0 & A \\ \overline{A} & 0 \end{pmatrix}$ 和 $\begin{pmatrix} 0 & B \\ \overline{B} & 0 \end{pmatrix}$ 相似，或者 $\begin{pmatrix} 0 & A \\ -\overline{A} & 0 \end{pmatrix}$ 和 $\begin{pmatrix} 0 & B \\ -\overline{B} & 0 \end{pmatrix}$ 相似，由此给出两个实方阵复相似的必要且充分条件.

2. 设

$$A\bar{A}=B\bar{B}=\text{diag}(N^{(e_1)},\cdots,N^{(e_t)}),e_1\geqslant\cdots\geqslant e_t>0.$$

什么时候A和B复相似？试用e_1, $\cdots$, e_t来表出A在复相似下的标准形$\text{diag}(N^{(f_1)},\cdots,N^{(f_s)})$中的数值$s$及$f_1,\cdots,f_s$.

3. 试给出n阶复方阵分解为Hermite方阵及复斜对称方阵的乘积的必要且充分条件.

4. 试求n阶复幂零方阵B能表成n阶复方阵A及$\bar{A}$之乘积之必要且充分条件.

第十四章　非 负 方 阵

§14.1　不可分拆非负方阵的特征根

定义　n 阶方阵称为**非负方阵**，如果它的元素全部为非负实数．n 阶非负方阵 A 称为**可分拆的**，如果存在置换方阵 P，使得

$$PAP' = \begin{pmatrix} A_1^{(s)} & A_2^{(s,n-s)} \\ 0 & A_3^{(n-s)} \end{pmatrix}$$

其中 s 有 $1 \leqslant s \leqslant n-1$．否则，称为**不可分拆的**．

为了讨论非负方阵，我们引进

定义　n 阶非负方阵 A 若非异，且逆方阵 A^{-1} 仍为非负方阵，则 A 称为**广义初等方阵**．

作为例子可知：置换方阵 P_{jk} 为广义初等方阵，又第二类初等方阵 $P_j(a)$，其中$a>0$，也都是广义初等方阵．但是第三类初等方阵不是广义初等方阵．

显然，非负方阵在广义初等方阵的相抵下仍为非负方阵．又广义初等方阵之乘积仍为广义初等方阵．

定理 14.1.1　广义初等方阵 P 能唯一地分解为

$$P = Q\Lambda_1 = \Lambda_2 Q,$$

其中 Q 为置换方阵，Λ_1，Λ_2 为对角元素大于零的对角方阵．因此 Q, Λ_1, Λ_2 也都是广义初等方阵．

证　记

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}, P^{-1} = \begin{pmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & & \vdots \\ q_{n1} & \cdots & q_{nn} \end{pmatrix}.$$

由于 P 为广义初等方阵，所以 p_{ij},q_{ij} 都是非负实数，今

$$\sum_{k=1}^{n} p_{ik}q_{kj}=\delta_{ij}, \qquad i,j=1,2,\cdots,n.$$

所以当 $i\neq j$，则有 $p_{ik}q_{kj}=0,k=1,2,\cdots,n$. 但是 P^{-1} 非异，所以它的第 k 行中必有一个元素 $q_{kj_0}\neq 0$. 因此 $p_{ik}=0,i=1,\cdots,j_0-1,j_0+1,\cdots,n$. 即 P 的第 k 列中恰有一个元素不等于零. 由 k 任取，所以证明了 P 的每一列恰有一个元素不等于零. 由 $\det P\neq 0$，所以 P 的每一行也恰有一个元素不等于零. 将 P 中非零元素换成 1，便得到一个置换方阵 Q，且 PQ 及 QP 都是对角方阵，分别记作 Λ_1，Λ_2. 它们的对角元素都大于零. 由 $Q^{-1}=Q$，这证明了存在如定理要求的分解式 $P=Q\Lambda_1=\Lambda_2 Q$. 由于 Q 为实正交方阵，Λ_i 为定正实对称方阵，所以上述分解实际上是极分解. 因此分解唯一. 证完.

引理 14.1.1 非负方阵的不可分拆性在广义初等方阵的相似下不改变.

证 设 A 为不可分拆非负方阵，P 为广义初等方阵. 由定理 14.1.1，$P=\Lambda_2 Q$，其中 Q 为置换方阵，Λ_2 为对角元素大于零的对角方阵. 于是

$$P^{-1}AP=Q(\Lambda_2^{-1}A\Lambda_2)Q.$$

由于 A 和 $\Lambda_2^{-1}A\Lambda_2$ 这两个非负方阵的同位置元素同时等于零或同时不等于零，这证明了 A 不可分拆当且仅当 $P^{-1}AP$ 不可分拆. 证完.

定理 14.1.2 设 A 为 n 阶非负方阵，则存在置换方阵 P 使得

$$P^{-1}AP=\begin{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1p}\\ & A_{22} & \cdots & A_{2p}\\ & & \ddots & \vdots \\ & & & A_{pp}\end{pmatrix} & \begin{pmatrix} A_{1,p+1} & \cdots & A_{1m}\\ A_{2,p+1} & \cdots & A_{2m}\\ \vdots & & \vdots\\ A_{p,p+1} & \cdots & A_{pm}\end{pmatrix} \\ 0 & \begin{pmatrix} A_{p+1,p+1} & & \\ & \ddots & \\ & & A_{mm}\end{pmatrix}\end{pmatrix},$$

其中 A_{11}，…，A_{mm} 为不可分拆非负方阵，且 $A_{j,j+1}$，…，A_{jm}，$1 \leqslant j \leqslant p$ 中必有一个非零矩阵. 又不计子块的次序，则标准形唯一.

证 如果 A 可分拆，则存在置换方阵 P_1，使得

$$P_1^{-1}AP_1=\begin{pmatrix} B_1^{(q)} & B_2 \\ 0 & B_3 \end{pmatrix},$$

对置换方阵 $P_2^{(q)}, P_3^{(n-q)}$，则

$$\begin{pmatrix} P_2 & 0 \\ 0 & P_3 \end{pmatrix}^{-1} P_1^{-1}AP_1\begin{pmatrix} P_2 & 0 \\ 0 & P_3 \end{pmatrix}=\begin{pmatrix} P_2^{-1}B_1P_2 & P_2^{-1}B_2P_3 \\ 0 & P_3^{-1}B_3P_3 \end{pmatrix}.$$

所以若 B_1, B_2 不可分拆，则已证明了定理. 若有一个可分拆，则变为

$$\begin{pmatrix} C_1 & C_2 & C_3 \\ 0 & C_4 & C_5 \\ 0 & 0 & C_6 \end{pmatrix}.$$

例如当 $C_2=0, C_3=0$，则作

$$\begin{pmatrix} 0 & I & 0 \\ I & 0 & 0 \\ 0 & 0 & I \end{pmatrix}^{-1}\begin{pmatrix} C_1 & 0 & 0 \\ 0 & C_4 & C_5 \\ 0 & 0 & C_6 \end{pmatrix}\begin{pmatrix} 0 & I & 0 \\ I & 0 & 0 \\ 0 & 0 & I \end{pmatrix}=\begin{pmatrix} C_4 & 0 & C_5 \\ 0 & C_1 & 0 \\ 0 & 0 & C_6 \end{pmatrix}.$$

按上面的思路，用归纳法便证明了标准形的存在性.

下面证唯一性. 为此用线性空间语言. 在 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \cdots, \alpha_n$. 于是存在 $\mathfrak{L}$ 上线性变换 $\mathscr{A}$，使得它在基 $\alpha_1, \cdots, \alpha_n$ 下的方阵表示为 A. 对方阵 A 用置换方阵作相似就等于对基的顺序 $\alpha_1\alpha_2\cdots\alpha_n$ 作排列. 所以用线性空间语言来描述这个定理，即要证明在 $\mathfrak{L}$ 中存在子空间 $\mathfrak{L}_1, \mathfrak{L}_2, \cdots, \mathfrak{L}_m$，使得 $\mathfrak{L}$ 有空间直接和

$$\mathfrak{L}=\mathfrak{L}_1\oplus\mathfrak{L}_2\oplus\cdots\oplus\mathfrak{L}_m,$$

且有

$$\mathscr{A}(\mathfrak{L}_1)\subset\mathfrak{L}_1, \mathscr{A}(\mathfrak{L}_2)\subset\mathfrak{L}_1+\mathfrak{L}_2, \cdots, \mathscr{A}(\mathfrak{L}_p)\subset\mathfrak{L}_1+\cdots+\mathfrak{L}_p,$$
$$\mathscr{A}(\mathfrak{L}_{p+1})\subset\mathfrak{L}_{p+1}+\mathfrak{L}_1+\cdots+\mathfrak{L}_p, \cdots,$$

$$\mathscr{A}(\mathfrak{L}_m)\subset\mathfrak{L}_m+\mathfrak{L}_1+\cdots+\mathfrak{L}_p.$$

而条件 A_{11}, …, A_{mm} 不可分拆等价于在 $\mathfrak{L}_j$ 中没有非零真子空间 $\mathfrak{M}_j$,使得$\mathscr{A}(\mathfrak{M}_j)\subset\mathfrak{M}_j+\mathfrak{L}_{j-1}+\cdots+\mathfrak{L}_1$, $j=1,2,\cdots,m$.

所谓有两种标准形,即存在 $\mathfrak{L}$ 的另一种子空间直接和分解

$$\mathfrak{L}=\mathfrak{S}_1\oplus\cdots\oplus\mathfrak{S}_q,$$

使得

$$\mathscr{A}(\mathfrak{S}_1)\subset\mathfrak{S}_1,\mathscr{A}(\mathfrak{S}_2)\subset\mathfrak{S}_1+\mathfrak{S}_2,\cdots,\mathscr{A}(\mathfrak{S}_s)\subset\mathfrak{S}_1+\mathfrak{S}_2+\cdots+\mathfrak{S}_s,$$

$$\mathscr{A}(\mathfrak{S}_{s+1})\subset\mathfrak{S}_{s+1}+\mathfrak{S}_1+\cdots+\mathfrak{S}_s,\cdots,\mathscr{A}(\mathfrak{S}_q)\subset\mathfrak{S}_q+\mathfrak{S}_1+\cdots+\mathfrak{S}_s,$$

且在 $\mathfrak{S}_j$ 中没有非零真子空间 $\mathfrak{R}_j$,使得

$$\mathscr{A}(\mathfrak{R}_j)\subset\mathfrak{R}_j+\mathfrak{S}_{j-1}+\cdots+\mathfrak{S}_1, j=1,2,\cdots,q.$$

为了证明唯一性,我们要证明 $q=m$,且 $\mathfrak{S}_1,\cdots,\mathfrak{S}_m$ 为 $\mathfrak{L}_1,\cdots,\mathfrak{L}_m$ 的一个排列.

注意到在$\alpha_1,\alpha_2,\cdots,\alpha_n$中可取一些元素使得它们是 $\mathfrak{L}_j$ 的基,且在$\alpha_1,\alpha_2,\cdots,\alpha_n$中可取一些元素使得它们是 $\mathfrak{S}_k$ 的基,这里 $j=1,2,\cdots,m$, $k=1,2,\cdots,q$. 所以存在 t 使得

$$\mathfrak{S}_1\cap\mathfrak{L}_1=0,\mathfrak{S}_1\cap\mathfrak{L}_2=0,\cdots,\mathfrak{S}_1\cap\mathfrak{L}_{t-1}=0,\mathfrak{S}_1\cap\mathfrak{L}_t\neq0.$$

今$\mathscr{A}(\mathfrak{S}_1)\subset\mathfrak{S}_1$, $\mathscr{A}(\mathfrak{L}_1+\cdots+\mathfrak{L}_t)\subset\mathfrak{L}_1+\cdots+\mathfrak{L}_t$. 所以由基底的选取可知

$$\mathscr{A}(\mathfrak{S}_1\cap\mathfrak{L}_t)\subset\mathfrak{S}_1\cap(\mathfrak{L}_1+\cdots+\mathfrak{L}_t)=\mathfrak{S}_1\cap\mathfrak{L}_t.$$

由条件 $\mathfrak{S}_1\cap\mathfrak{L}_t\neq0$, 又 $\mathfrak{S}_1$ 中无$\mathscr{A}$之不变子空间,所以证明了

$$\mathfrak{S}_1\cap\mathfrak{L}_t=\mathfrak{S}_1$$

即 $\mathfrak{S}_1\subset\mathfrak{L}_t$. 这证明了在 $\mathfrak{L}_t$ 中有非零子空间 $\mathfrak{S}_1$,使得

$$\mathscr{A}(\mathfrak{S}_1)\subset\mathfrak{S}_1+\mathfrak{L}_{t-1}+\cdots+\mathfrak{L}_1.$$

由条件便证明了 $\mathfrak{S}_1=\mathfrak{L}_t$.

由于在 $\mathfrak{S}_1,\cdots,\mathfrak{S}_q$ 中各存在基, 它们拼成$\{\alpha_1,\cdots,\alpha_n\}$, 在 $\mathfrak{L}_1$, …, $\mathfrak{L}_m$ 中也各存在基,它们也拼成$\{\alpha_r,\cdots,\alpha_n\}$. 所以证明了

$$\mathfrak{S}_2+\cdots+\mathfrak{S}_q=\mathfrak{L}_1+\cdots+\mathfrak{L}_{t-1}+\mathfrak{L}_{t+1}+\cdots+\mathfrak{L}_m.$$

另一方面，由于 $\mathfrak{L}_t$ 为$\mathscr{A}$的不变子空间，所以

$$\mathfrak{L}=\mathfrak{L}_t\oplus\mathfrak{L}_1\oplus\cdots\oplus\mathfrak{L}_{t-1}\oplus\mathfrak{L}_{t+1}\oplus\cdots\oplus\mathfrak{L}_m$$

仍适合同样条件. 所以无妨设 $\mathfrak{S}_1=\mathfrak{L}_1$.

考虑商空间$\mathfrak{L}/\mathfrak{L}_1$(定义见附录2的最后). 由于$\mathscr{A}(\mathfrak{L}_1)\subset\mathfrak{L}_1$，所以$\mathscr{A}$诱导了$\mathfrak{L}/\mathfrak{L}_1$上线性变换$\widetilde{\mathscr{A}}$，且有

$$\mathfrak{L}/\mathfrak{L}_1=\mathfrak{L}_2/\mathfrak{L}_1\oplus\cdots\oplus\mathfrak{L}_p/\mathfrak{L}_1\oplus\mathfrak{L}_{p+1}/\mathfrak{L}_1\oplus\cdots\oplus\mathfrak{L}_m/\mathfrak{L}_1,$$

$$\mathfrak{L}/\mathfrak{L}_1=\mathfrak{L}/\mathfrak{S}_1=\mathfrak{S}_2/\mathfrak{S}_1\oplus\cdots\oplus\mathfrak{S}_s/\mathfrak{S}_1\oplus\mathfrak{S}_{s+1}/\mathfrak{S}_1\oplus\cdots\oplus\mathfrak{S}_8/\mathfrak{S}_1.$$

对线性变换$\widetilde{\mathscr{A}}$而言，它们仍然适合定理条件. 且$\mathfrak{L}_2/\mathfrak{L}_1$, $\cdots$, $\mathfrak{L}_m/\mathfrak{L}_1$ 中有基，它们构成集合$\{\alpha_1+\mathfrak{L}_1, \cdots, \alpha_n+\mathfrak{L}_1\}-\{0+\mathfrak{L}_1\}$；同理，$\mathfrak{S}_2/\mathfrak{S}_1$, $\cdots$, $\mathfrak{S}_q/\mathfrak{S}_1$ 中有基，它们构成集合$\{\alpha_1+\mathfrak{L}_1, \cdots, \alpha_n+q_1\}-\{0+\mathfrak{L}_1\}$. 所以对维数作归纳法，便证明了在差一个

排列意义下，无妨设 $\mathfrak{S}_j/\mathfrak{S}_1=\mathfrak{L}_j/\mathfrak{L}_1$, $j=1,2,\cdots,m$, 且 $q=m$. 由于 $\mathfrak{S}_1=\mathfrak{L}_1$, 且 $\mathfrak{S}_1,\cdots,\mathfrak{S}_m$ 中有基拼成$\{\alpha_1, \cdots,\alpha_n\}$, $\mathfrak{L}_1,\cdots,\mathfrak{L}_m$ 中有基拼成$\{\alpha_1,\cdots,\alpha_n\}$，这证明了

$$q=m, \mathfrak{S}_j=\mathfrak{L}_j, j=1,2,\cdots,m.$$

定理证完.

给定 n 阶非负方阵 A，定理 14.1.2 给出了标准形. 于是有

$$\det(\lambda I-A)=\det(\lambda I-P^{-1}AP)=\prod_{j=1}^{m}\det(\lambda I-A_j).$$

因此求非负方阵的特征根的问题化为求不可分拆非负方阵的特征根. 下面来讨论不可分拆非负方阵的特征根

定理 14.1.3 设 A 为非零方阵，且为不可分拆非负方阵. 则 A 有正特征根 ρ 及属于 ρ 之特征向量 $(d_1,\cdots,d_n)'$, 其中 $d_1>0$, $\cdots,d_n>0$. 且若另有正特征根 ρ_0 及相应特征向量$(f_1,\cdots,f_n)'$，则有 $\rho_0=\rho$，且$f_j=ad_j$, $1\leqslant j\leqslant n$，其中 a 为正实数.

证 取定 n 阶非负方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\neq 0.$$

记

$$\rho(A)=\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}a_{ij}>0,$$

$$\rho=\operatorname*{Inf}_{\substack{\text{对一切广义}\\ \text{初等方阵}P}}\rho(P^{-1}AP)\geqslant 0.$$

我们来证 $\rho>0$，且 ρ 为 A 的正特征根，并求出一个属于特征根 ρ 的特征向量 $(d_1,\cdots,d_n)'$，其中 $d_1>0,\cdots,d_n>0$.

今任取广义初等方阵 P，由定理 14.1.1，$P=\Lambda Q$，其中 Q 为置换方阵，Λ 为对角元素大于零的对角方阵，记作

$$\Lambda=\mathrm{diag}(d_1,\cdots,d_n).$$

于是

$$\rho(P^{-1}AP)=\rho(Q^{-1}\Lambda^{-1}A\Lambda Q).$$

记由 Q 决定的 $1,2,\cdots,n$ 的排列为 $i_1i_2\cdots i_n$，则

$$Q^{-1}\Lambda^{-1}A\Lambda Q=\begin{pmatrix} a_{i_1i_1}d_{i_1}^{-1}d_{i_1} & \cdots & a_{i_1i_n}d_{i_1}^{-1}d_{i_n} \\ \vdots & & \vdots \\ a_{i_ni_1}d_{i_n}^{-1}d_{i_1} & \cdots & a_{i_ni_n}d_{i_n}^{-1}d_{i_n} \end{pmatrix}.$$

于是

$$\rho(P^{-1}AP)=\max_{1\leqslant k\leqslant n}\sum_{l=1}^{n}a_{i_ki_l}d_{i_k}^{-1}d_{i_l}=\max_{1\leqslant i\leqslant n}a_{ij}d_i^{-1}d_j=\rho(\Lambda^{-1}A\Lambda).$$

因此

$$\rho=\operatorname*{Inf}_{\substack{\Lambda=\mathrm{diag}(d_1,\cdots,d_n)\\ d_1>0,\cdots,d_n>0}}\rho(\Lambda^{-1}A\Lambda)=\operatorname*{Inf}_{\substack{\Lambda=\mathrm{diag}(d_1,\cdots,d_n)\\ d_1>0,\cdots,d_n>0,\ \sum_{j=1}^{n}d_j=1}}\rho(\Lambda^{-1}A\Lambda).$$

由下确界的定义，所以存在

$$\Lambda_k=\mathrm{diag}(d_{1k},\cdots,d_{nk}),\quad \wedge d_k0,\cdots,d_{nk}>0,\quad \sum_{j=1}^{n}d_{jk}=1,$$

使得

$$\lim\rho(\Lambda_k^{-1}A\Lambda_k)=\rho$$

今$\{d_{i1},d_{i2},\cdots\}$为正项有界序列，所以有收敛子序列，$i=1,2,\cdots,n$. 因此无妨设每个序列$\{d_{i1},d_{i2},\cdots\}$收敛于d_i, $i=1,2,\cdots,n$. 自然$d_1\geqslant0,\cdots,d_n\geqslant0$.

今对A的行及列同步地作对换，它诱导了$d_1,\cdots,d_n$作排列. 所以无妨设

$$d_1\geqslant\cdots\geqslant d_r>d_{r+1}=\cdots=d_n=0,\sum_{j=1}^{r}d_j=1.$$

又记

$$\Lambda_0=\mathrm{diag}(d_1,\cdots,d_n).$$

则有$\lim\limits_{k\to\infty}\Lambda_k=\Lambda_0$, 且当$k$充分大，有

$$\rho\leqslant\rho(\Lambda_k^{-1}A\Lambda_k)<\rho+\frac{1}{k}.$$

我们先来证明$r=n$. 设若$r<n$, 按前r行列分块，有

$$A=\begin{pmatrix}A_1^{(r)} & A_2\\ A_3 & A_4\end{pmatrix},\ \Lambda_k=\begin{pmatrix}D_k & 0\\ 0 & E_k\end{pmatrix},\ \Lambda_0=\begin{pmatrix}D_0 & 0\\ 0 & 0\end{pmatrix}.$$

因此$D_k\to D_0, E_k\to0$. 又

$$\Lambda_k^{-1}A\Lambda_k=\begin{pmatrix}D_k^{-1}A_1D_k & D_k^{-1}A_2E_k\\ E_k^{-1}A_3D_k & E_k^{-1}A_4E_k\end{pmatrix}.$$

今 $\rho(\Lambda_k^{-1}A\Lambda_k)<\rho+\frac{1}{k}$, 所以$\Lambda_k^{-1}A\Lambda_k$的各行之和都小于$\rho+\frac{1}{k}$. 特别，$E_k^{-1}A_3D_k$的各行之和都小于$\rho+\frac{1}{k}$. 这证明了

$$\sum_{j=1}^{r}a_{ij}d_{ik}^{-1}d_{jk}<\rho+\frac{1}{k},\ i=r+1,\cdots,n.$$

即

$$\sum_{j=1}^{r} a_{ij}d_{jk} < \left(\rho+\frac{1}{k}\right)d_{ik},\ i=r+1,\cdots,n.$$

当 $k\to\infty$，便有 $\sum_{j=1}^{r} a_{ij}d_j=0$，所以 $a_{ij}d_j=0, r+1\leqslant i\leqslant n,\ 1\leqslant j\leqslant r$. 由 $d_1>0,\cdots,d_r>0$，便证明了 $a_{ij}=0,\ 1\leqslant j\leqslant r,\ r+1\leqslant i\leqslant n$，即 $A_3=0$. 这和 A 不可分拆的假设矛盾. 所以证明了 $r=n$.

于是 $\Lambda_k\to\Lambda_0, \Lambda_k^{-1}\to\Lambda_0^{-1}$. 因此由 $\rho\leqslant\rho(\Lambda_k^{-1}A\Lambda_k)<\rho+\frac{1}{k}$，则有 $\rho(\Lambda_0^{-1}A\Lambda_0)=\rho>0$.

记 $B=\Lambda_0^{-1}A\Lambda_0=(b_{ij})$，于是 $b_{ij}=a_{ij}d_i^{-1}d_j,\ 1\leqslant i_{ij},\leqslant n$. 因此令

$$\rho_i=\sum_{j=1}^{n} b_{ij}=\sum_{j=1}^{n} a_{ij}d_i^{-1}d_j, 1\leqslant i\leqslant n.$$

则 $\rho=\rho(\Lambda_0^{-1}A\Lambda_0)=\max\limits_{1\leqslant i\leqslant n}\rho_i$. 我们无妨设

$$\rho=\rho_1=\cdots=\rho_s>\rho_{s+1}\geqslant\cdots\geqslant\rho_n>0.$$

下面来证明 $s=n$. 设若不然，即 $s<n$. 将 B 按前 s 行列分块为

$$B=\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}.$$

由于 A 不可分拆，所以 B 也不可分拆，即有 $B_2\neq 0$、将 B 的前 s 行及列，后 $n-s$ 行及列作同步的置换，从而无妨设 B_2 的第 1 行第 n 列元素 $b_{1n}\neq 0$. 作

$$C_1=P_n(a)^{-1}BP_n(a)=\begin{pmatrix} b_{11} & \cdots & b_{1,n-1} & ab_{1n} \\ \vdots & & \vdots & \vdots \\ b_{n-,11} & \cdots & b_{n-1,n-1} & ab_{n-1,n} \\ a^{-1}b_{n1} & \cdots & a^{-1}b_{n,n-1} & b_{nn} \end{pmatrix},$$

于是各行和分别为

$$\rho_i' = \rho_i + (a-1)b_{in}, \quad i=1,2,\cdots,n-1,$$
$$\rho_n' = a^{-1}\rho_n + (1-a^{-1})b_{nn}.$$

当 $0<a<1$ 时便有 $\rho_1'<\rho_1, \rho_i'\leqslant\rho_i$, $2\leqslant i<n$. 由 $b_{1n}>0$, 所以可取 a 使得

$$0<a<1, \quad a^{-1}\rho_n + (1-a^{-1})b_{nn} < \rho_1 + (a-1)b_{1n}.$$

此即 $\rho_n'<\rho_1'$. 所以

$$\rho_1'<\rho_1, \rho_i'\leqslant\rho_i, 2\leqslant i\leqslant n-1, \rho_n'<\rho_1'<\rho_1.$$

于是由 $\rho(B)=\mathrm{Inf}\,\rho(P^{-1}AP)$, 便有

$$\rho_1 = \max_{1\leqslant i\leqslant n}\rho_i \geqslant \max_{1\leqslant i\leqslant n}\rho_i' = \rho(C_1)\geqslant\rho(B)=\rho_1$$

这证明了存在 $\rho_{i_0}'=\rho_1$, 且 $\rho(C_1)=\rho(B)$, 这里 $i_0\in\{2,\cdots,s\}$, 这证明了至多 $s-1$ 个 ρ_j' 等于 ρ.

对 C_1 用上法讨论, 所以至多 s 步, 便得到非负方阵序列 C_1, $C_2,\cdots,C_q$, 使得它们都由对 B 作广义初等方阵之相似所得到, 且 $\rho=\rho(B)=\rho(C_1)=\cdots=\rho(C_q)$, 但 C_q 的每个行和小于 ρ. 这和 $\rho=\mathrm{Inf}\,\rho(P^{-1}AP)$ 矛盾. 所以证明了 $s=n$, 即有

$$\rho=\rho_1=\cdots=\rho_n.$$

因此

$$A\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \begin{pmatrix} \Sigma a_{1j}d_j \\ \vdots \\ \Sigma a_{nj}d_j \end{pmatrix} = \begin{pmatrix} d_1\rho_1 \\ \vdots \\ d_n\rho_n \end{pmatrix} = \rho\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$$

所以 ρ 为 A 的正特征根, 而 $(d_1, \cdots, d_n)'$ 为 A 的属于特征根 ρ 之特征向量, 其中 $d_1>0,\cdots,d_n>0$.

下面证明唯一性. 若 A 有一个正特征根 ρ_0, 且 A 的属于特征根 ρ_0 之特征向量为 $(f_1,\cdots,f_n)_1'$ 使得 $f_1>0,\cdots,f_n>0$. 我们来证 $\rho_0=\rho$, 且存在正实数 a, 使得 $f_j=ad_j$, $j=1,2,\cdots,n$. 事实上, 令

$$D=\mathrm{diag}(d_1,\cdots,d_n), \quad F=\mathrm{diag}(f_1,\cdots,f_n).$$
$$B=D^{-1}AD=(b_{ij}), \quad C=F^{-1}AF=(C_{ij}).$$

记 $e=(1,1,\cdots,1)'$. 则由 $De=(d_1,\cdots,d_n)'$, $Fe=(f_1,\cdots,f_n)'$, 所以 $ADe=\rho De$, $AFe=\rho_0 Fe$, 此即

$$Be=\rho e,\quad Ce=\rho_0 e.$$

但是 $A=DBD^{-1}=FCF^{-1}$, 即 $C=(F^{-1}D)B(F^{-1}D)^{-1}$, 其中

$$\Lambda=F^{-1}D=\operatorname{diag}(f_1^{-1}d_1,\cdots,f_n^{-1}d_n).$$

记 $\lambda_j^{-1}=f_j^{-1}d_j$, $1\leqslant j\leqslant n$. 显然可无妨设

$$\lambda_1=\cdots=\lambda_t>\lambda_{t+1}\geqslant\cdots\geqslant\lambda_n>0.$$

今 $(F^{-1}D)^{-1}C=B(F^{-1}D)^{-1}$, 所以有 $\lambda_i c_{ij}=b_{ij}\lambda_j$. 由 $Be=\rho e$, $Ce=\rho_0 e$, 所以有 $\sum\limits_j b_{ij}=\rho$, $\sum\limits_j c_{ij}=\rho_0$. 因此

$$\rho_0=\sum_{j=1}^n c_{ij}=\sum_{j=1}^n \lambda_i^{-}\lambda_j b_{ij},\quad i=1,2,\cdots,n.$$

于是

$$\rho\min_{1\leqslant k\leqslant n}\lambda_k\leqslant\sum_{j=1}^n b_{ij}\lambda_j\leqslant\rho\max_{1\leqslant k\leqslant n}\lambda_k,\quad 1\leqslant i\leqslant n.$$

因此由 $\sum\limits_{j=1}^n b_{ij}\lambda_j=\lambda_i\rho_0$, 便得

$$\lambda_n\rho\leqslant\lambda_i\rho_0\leqslant\lambda_1\rho_j, 1\leqslant i\leqslant n.$$

取 $i=n$, 有 $\rho\leqslant\rho_0$, 取 $i=1$, 有 $\rho_0\leqslant\rho$. 这证明了 $\rho_0=\rho$.
因此有

$$\sum_{j=1}^n b_{ij}\lambda_j=\lambda_i\rho_0=\lambda_i\rho=\sum_{j=1}^n\lambda_i b_{ij},$$

即

$$\sum_{j=1}^n(\lambda_i-\lambda_j)b_{ij}=0,\quad 1\leqslant j\leqslant n.$$

但是 $\lambda_1=\cdots=\lambda_t>\lambda_{t+1}\geqslant\cdots\geqslant\lambda_n>0$, 取 $i=1,2,\cdots,t$, 便证明了 $\sum\limits_{j=t+1}^n(\lambda_1-\lambda_j)b_{ij}=0$, 其中 $\lambda_1-\lambda_j>0$, 所以 $b_{ij}=0$, $1\leqslant i\leqslant t$,

$t+1\leqslant j\leqslant n$，即证明了当 $n>t$ 时 B 可分拆，这和假设 B 不可分拆矛盾，所以证明了 $n=t$，即$\Lambda=F^{-1}D=\lambda_1^{-1}I^{(n)}$，即 $D=\lambda_1^{-1}F$，所以 $f_j=\lambda_1 d_j$，$1\leqslant j\leqslant n$，其中 λ_1 为正实数，定理证完．

定理 14.1.4(比较定理) 设 A 为 n 阶不可分拆非负方阵，且

$$Ae=\rho e.$$

设 B 为 n 阶复方阵，记 $B=(b_{ij})$，$A=(a_{ij})$．假设

$$|b_{ij}|\leqslant a_{ij},\quad 1\leqslant i,j\leqslant n.$$

则 B 的任一特征根 λ_0 适合条件

$$|\lambda_0|\leqslant\rho,$$

且 B 有形如 $\lambda_0=\rho e^{\sqrt{-1}\theta}$的特征根当且仅当 λ_0 的特征向量为 $[\exp(-\sqrt{-1}\Lambda)]e$，又

$B=e^{\sqrt{-1}}\theta[\exp(-\sqrt{-1}\Lambda)]A[\exp(\sqrt{-1}\Lambda)]$，$\Lambda=\mathrm{diag}(\theta_1,\cdots,\theta_n)$，其中 $0\leqslant\theta_1\leqslant\cdots\leqslant\theta_n<2\pi$．

证 设 λ_0 为 B 的特征根，$\alpha=(u_1,u_2,\cdots,u_n)'$ 为相应的特征向量．由于 $B\alpha=\lambda_0\alpha$，即有

$$\sum_{j=1}^n b_{kj}u_j=\lambda_0 u_k,\quad 1\leqslant k\leqslant n.$$

记$u_0=\max\limits_{1\leqslant k\leqslant n}|u_k|$，于是 $0\neq u_0=|u_{k_0}|$．且由$|b_{ij}|\leqslant a_{ij}$，有

$$|\lambda_0|u_0=|\lambda_0 u_{k_0}|=\left|\sum_{j=1}^n b_{k_0j}u_j\right|\leqslant\sum_{j=1}^n|b_{k_0j}|u_0\leqslant\sum_{j=1}^n a_{k_0j}u_0=\rho u_0.$$

因此证明了$|\lambda_0|\leqslant\rho$．

设 B 有特征根 $\lambda_0=\rho e^{\sqrt{-1}\theta}$．显然无妨设它的特征向量 α 有

$$|u_1|=\cdots=|u_t|>|u_{t+1}|\geqslant\cdots\geqslant|u_n|\geqslant 0.$$

我们来证 $t=n$．事实上，若$t<n$，将 A 按前 t 行列分块为

$$A=\begin{pmatrix}A_1 & A_2\\ A_3 & A_4\end{pmatrix}.$$

由 A 不可分拆，$1\leqslant t<n$，所以 $A_2\neq 0$．对 A 的前 t行、列，对A的后

$n-t$ 行、列分别作同步的置换，于是无妨设 $a_{1n}\neq 0$. 而

$$\rho|u_1|=|\rho e^{\sqrt{-1}\theta}u_1|=|\lambda_0 u_1|=\left|\sum_{j=1}^{n}b_{1j}u_j\right|\leqslant\sum_{j=1}^{n}a_{1j}|u_j|.$$

由于 $a_{1n}|u_n|<a_{1n}|u_1|$，所以

$$\rho|u_1|<\sum_{j=1}^{n}a_{1j}|u_1|=\rho|u_1|.$$

这导出矛盾. 因此证明了 $t=n$，即有

$$|u_1|=\cdots=|u_n|=a>0,$$

即

$$u_j=ae^{-\sqrt{-1}\theta_j},\quad j=1,2,\cdots,n.$$

从特征向量角度可知无妨设 $a=1$. 所以特征向量

$$\alpha=\begin{pmatrix}u_1\\ \vdots\\ u_n\end{pmatrix}=\begin{pmatrix}e^{-\sqrt{-1}\theta_1}\\ \vdots\\ e^{-\sqrt{-1}\theta_n}\end{pmatrix}=\begin{pmatrix}e^{-\sqrt{-1}\theta_1} & & \\ & \ddots & \\ & & e^{-\sqrt{-1}\theta_n}\end{pmatrix}$$

$$e=(\exp(-\sqrt{-1}\Lambda))e,$$

其中 $\Lambda=\mathrm{diag}(\theta_1,\ \cdots,\theta_n)$. 显然无妨设 $0\leqslant\theta_1\leqslant\cdots\leqslant\theta_n<2\pi$.

现在来计算 n 阶复方阵 B. 今

$$B\alpha=\lambda_0\alpha=\rho e^{\sqrt{-1}\theta}e^{-\sqrt{-1}\Lambda}e,$$

故有 $Be^{-\sqrt{-1}\Lambda}e=\rho e^{\sqrt{-1}(\theta I-\Lambda)}e$，即有

$$\sum_{j=1}^{n}b_{ij}e^{-\sqrt{-1}\theta_j}=\rho e^{\sqrt{-1}(\theta-\theta_i)},\ i=1,2,\cdots,n.$$

此即

$$\rho=\sum_{j=1}^{n}b_{ij}e^{\sqrt{-1}(\theta_i-\theta-\theta_j)}=\left|\sum_{j=1}^{n}b_{ij}e^{\sqrt{-1}(\theta_i-\theta-\theta_j)}\right|$$

$$\leqslant\sum_{j=1}^{n}|b_{ij}|\leqslant\sum_{j=1}^{n}a_{ij}$$

而 $Ae=\rho e$，即有 $\rho=\sum\limits_{j=1}^{n}a_{ij}$，这证明了

$$|b_{ij}|=a_{ij},\quad i,j=1,2,\cdots,n.$$

所以

$$b_{ij}=a_{ij}e^{\sqrt{-1}\xi_{ij}},\qquad i,j=1,2,\cdots,n.$$

于是

$$\rho=\sum_{j=1}^{n}a_{ij}=\sum_{j=1}^{n}b_{ij}e^{\sqrt{-1}(\theta_i-\theta-\theta_j)}=\sum_{j=1}^{n}a_{ij}e^{\sqrt{-1}(\xi_{ij}+\theta_i-\theta-\theta_j)},$$

$$j=1,2,\cdots,n.$$

所以

$$\sum_{j=1}^{n}a_{ij}=\sum_{j=1}^{n}a_{ij}\cos(\xi_{ij}+\theta_i-\theta-\theta_j)$$

$$=\sum_{j=1}^{n}a_{ij}-2\sum_{j=1}^{n}a_{ij}\sin^2\frac{1}{2}(\xi_{ij}+\theta_i-\theta-\theta_j).$$

这证明了当 $a_{ij}>0$，有 $\sin\frac{1}{2}(\xi_{ij}+\theta_i-\theta-\theta_j)=0$，即有

$$\xi_{ij}=-\theta_i+\theta+\theta_j+2k_{ij}\pi,$$

其中 k_{ij} 为整数． 在 $a_{ij}=0$ 时，约定 $\xi_{ij}=-\theta_i+\theta+\theta_j$． 总之，有

$$e^{\sqrt{-1}\xi_{ij}}=e^{\sqrt{-1}(\theta+\theta_j-\theta_i)},i,j=1,2,\cdots,n.$$

于是

$$B=(b_{ij})=(a_{ij}e^{\sqrt{-1}\xi_{ij}})=(a_{ij}e^{\sqrt{-1}(\theta+\theta_j-\theta_i)})$$

$$=e^{\sqrt{-1}\theta}(\exp\sqrt{-1}\Lambda)^{-1}A\exp\sqrt{-1}\Lambda.$$

反之，若$B=e^{\sqrt{-1}\theta}(\exp\sqrt{-1}\Lambda)^{-1}A\exp\sqrt{-1}\Lambda)$有特征根$\lambda_0$，相应特征向量为$(\exp\sqrt{-1}\Lambda)^{-1}e$，我们来证 $\lambda_0=\rho e^{\sqrt{-1}\theta}$．事实上，由 $B(\exp\sqrt{-1}\Lambda)^{-1}e=\lambda_0(\exp\sqrt{-1}\Lambda)^{-1}e$，所以

$$\lambda_0e=(\exp\sqrt{-1}\Lambda)B(\exp\sqrt{-1})^{-1}e=e^{\sqrt{-1}\theta}Ae=\rho e^{\sqrt{-1}\theta}e.$$

这证明了 $\lambda_0=\rho e^{\sqrt{-1}\theta}$． 定理证完．

定理 14.1.5 n 阶不可分拆非负方阵 A 的任一特征根 λ_0 有

$$|\lambda_0| \leqslant \rho = \operatorname*{Inf}_{\substack{\text{对一切广义} \\ \text{初等方阵} P}} \rho(P^{-1}AP),$$

且 A 的正特征根 ρ 为单重根. 又 $\rho I^{(n)} - A$ 的伴随方阵 $A(\rho)$ 的元素都是正实数. 记 $A = (a_{ij})$, 则有

$$\min_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij} \leqslant \rho \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij},$$

且有

$$\rho = \min_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij} \text{ 当且仅当 } \rho = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij}.$$

证 将 A 的第 k 行及第 k 列元素全换成零，得到的非负方阵记作 A_k, 将 A_k 的第 k 行及第 k 列元素除去, 得到 A 的子矩阵记作 B_k, $k = 1, 2, \cdots, n$. 于是

$$\frac{d}{d\lambda} \det(\lambda I^{(n)} - A) = \sum_{k=1}^{n} \det(\lambda I^{(n-1)} - B_k).$$

由定理 14.1.3 A 有特征根 ρ 及属于特征根 ρ 的特征向量 $(d_1, \cdots, d_n)'$. 记 $D = \operatorname{diag}(d_1, \cdots, d_n)$, 则由

$$A \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \rho \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix},$$

有

$$ADe = \rho De.$$

记 $\tilde{A} = D^{-1}AD$, 则 $\tilde{A}$ 仍为不可分拆非负方阵. 且

$$\tilde{A}e = \rho e.$$

在定理 14.1.4 中取 A_k 为 B, 由 A_k 的定义可知定理条件适合. 因此证明了 A_k 的特征根 λ_0 有 $|\lambda_0| \leqslant \rho$, 且若 A_k 有特征根 $\lambda_0 = \rho e^{i\theta}$, 对应的特征向量为 $[\exp(-\sqrt{-1}\Lambda)]e$, 又

$$A_k = e^{\sqrt{-1}\theta} [\exp(-\sqrt{-1}\Lambda)] A [\exp(\sqrt{-1}\Lambda)],$$

这里 $\Lambda=\mathrm{diag}(\theta_1,\cdots,\theta_n)$，但是由于 A_k 的第 k 行和第 k 列全等于零，所以 A_k 可分拆，这证明了 A 可分拆，因此导出矛盾. 所以 $\tilde{A}_k=D^{-1}A_kD$ 的所有特征根之模小于 ρ.

在 $\tilde{A}_k$ 中划去第 k 行和第 k 列，记作 $\tilde{B}_k$. 由

$$\det(\lambda I^{(n)}-\tilde{A}_k)=\lambda\det(\lambda I^{(n-1)}-\tilde{B}_k).$$

取 $\lambda=\rho$，则由 $\det(\rho I^{(n)}-\tilde{A}_k)>0$，所以 $\det(\rho I^{(n-1)}-\tilde{B}_k)>0$.

由 $\dfrac{d}{d\lambda}\det(\lambda I^{(n)}-A)=\dfrac{d}{d\lambda}\det(\lambda I^{(n)}-\tilde{A})=\sum\limits_{k=1}^{n}\det(\lambda I^{(n-1)}-\tilde{B}_k)$.

这证明了

$$\frac{d}{d\lambda}\det(\lambda I^{(n)}-A)\big|_{\lambda=\rho}>0.$$

所以 $\det(\lambda I^{(n)}-A)$ 以 ρ 为单重根.

今 $\det(\rho I^{(n)}-A)=0$. 对 $\rho I^{(n)}-A$ 的伴随方阵 $A(\rho)$，由 Laplace 展开定理可知

$$(\rho I-A)A(\rho)=\det(\rho I-A)\cdot I=0.$$

记 $e_j\in R^n$，其中 e_j 的第 j 个坐标为1，其余坐标为零，又记 $\alpha_j=A(\rho)e_j$，即为 $A(\rho)$ 的第 j 列. 于是有

$$A\alpha_j=\rho\alpha_j,\ j=1,2,\cdots,n.$$

由于 A 以 ρ 为单重特征根，即特征向量生成一维子空间，所以对 A 的属于 ρ 的特征向量 $\alpha=(d_1,\cdots,d_n)'$, $d_1>0,\cdots,d_n>0$，则有

$$A(\rho)e_j=\alpha_j=a_j\alpha,\ j=1,2,\cdots,n.$$

其中 $a_1,\cdots,a_n$ 为实数. 所以

$$A(\rho)=(a_1\alpha\cdots a_n\alpha).$$

下面证 $A(\rho)$ 的元素全为正实数，即证 $a_1>0,\cdots,a_n>0$. 事实上由于 $\det(\rho I-B_k)$ 为 $A(\rho)$ 的第 k 个对角元素，而已证 $\det(\rho I^{(n-1)}-\tilde{B}_k)>0$，又 $\tilde{B}_k$ 和 B_k 相似，所以证明了 $A(\rho)$ 的第 k 个对角元素 $a_kd_k>0$, $k=1,2,\cdots,n$ 已知 $d_1>0,\cdots,d_n>0$，这证明了 $a_1>0,\cdots,$

$a_n>0$. 所以 $A(\rho)$ 由正实数构成.

记 $A=(a_{ij})$, $\rho_i=\sum_{j=1}^{n}a_{ij}$, $i=1,2,\cdots,n$. 下面证

$$\min_{1\leqslant i\leqslant n}\rho_i\leqslant\rho\leqslant\max_{1\leqslant i\leqslant n}\rho_i.$$

由 $\rho=\operatorname{Inf}\rho(P^{-1}AP)\leqslant\rho(A)=\max_{1\leqslant i\leqslant n}\rho_i$, 所以只要证 $\min_{1\leqslant i\leqslant n}\rho_i\leqslant\rho$. 今

$$0=\det(\rho I-A)=\det\begin{pmatrix}\rho-a_{11} & \cdots & -a_{1n}\\ \vdots & & \vdots\\ -a_{n1} & \cdots & \rho-a_{nn}\end{pmatrix}$$

$$=\det\begin{pmatrix}\rho-a_{11} & \cdots & -a_{1,n-1} & \rho-\rho_1\\ \vdots & & \vdots & \vdots\\ -a_{n-1,1} & \cdots & \rho-a_{n-1,n-1} & \rho-\rho_{n-1}\\ -a_{n1} & \cdots & -a_{n,n-1} & \rho-\rho_n\end{pmatrix}$$

按最后一列作 Laplace 展开,则有

$$\sum_{j=1}^{n}(\rho-\rho_j)A_{jn}(\rho)=0,$$

其中 $A_{jn}(\rho)$ 为 $A(\rho)$ 的第 n 行第 j 列元素,所以都是正实数,这证明了 $\rho-\rho_1,\cdots,\rho-\rho_n$ 中若有元素大于零,则也有元素小于零,反之亦然,所以 $\min_{1\leqslant i\leqslant n}\rho_i\leqslant\rho\leqslant\max_{1\leqslant i\leqslant n}\rho_i$.

最后,设 $\rho=\min_{1\leqslant i\leqslant n}\rho_i$,则 $\rho\leqslant\rho_i$, $1\leqslant i\leqslant n$, 所以 $\rho-\rho_i\leqslant0$, $i=1,2,\cdots,n$. 由 $\sum_{j=1}^{n}(\rho-\rho_j)A_{jn}(\rho)=0$ 可知 $\rho_i=\rho$, $1\leqslant i\leqslant n$. 因此 $\rho=\max_{1\leqslant i\leqslant n}\rho_i$. 同理,设 $\rho=\max_{1\leqslant i\leqslant n}\rho_i$,则 $\rho\geqslant\rho_i$, $1\leqslant i\leqslant n$,所以 $\rho_i=\rho$, $1\leqslant i\leqslant n$. 因此 $\rho=\min_{1\leqslant i\leqslant n}\rho_i$. 定理证完.

定理 14.1.6 设 A 为 n 阶不可分拆非负方阵，ρ 为 A 的最大正特征根，则模等于 ρ 的特征根都是单重特征根，它们是

$$\varepsilon_1\rho, \varepsilon_2\rho, \cdots, \varepsilon_m\rho,$$

其中 m 为自然数，$\varepsilon_j = \exp\frac{2\sqrt{-1}(j-1)\pi}{m}, j=1, 2, \cdots, m$. 它们对应的特征向量分别为

$$(\exp\sqrt{-1}\Lambda_1)^{-1}\alpha,\ (\exp\sqrt{-1}\Lambda_2)^{-1}\alpha, \cdots, (\exp\sqrt{-1}\Lambda_m)^{-1}\alpha,$$

其中 $\Lambda_j = \text{diag}(\theta_{j1}, \cdots, \theta_{jn})$，$0 \leqslant \theta_{j1}, \cdots, \theta_{jn} < 2\pi$，$j=1,2,\cdots,m$. 又 $\alpha = (d_1, \cdots, d_n)', d_1 > 0, \cdots, d_n > 0$. 再

$$A = e^{\frac{2\sqrt{-1}\pi}{m}}(\exp\sqrt{-1}\Lambda_1)^{-1}A\exp\sqrt{-1}\Lambda_1.$$

另一方面，存在置换方阵 P，使得

$$P^{-1}e^{\sqrt{-1}\Lambda_1}P = \text{diag}(\varepsilon_1 I^{(n_1)}, \cdots, \varepsilon_m I^{(n_m)}),$$

$$P^{-1}AP = \begin{pmatrix} 0 & \cdots & 0 & A_{1m} \\ A_{21} & \ddots & \vdots & \vdots \\ & \ddots & 0 & 0 \\ & & A_{m,m-1} & 0 \end{pmatrix}.$$

证 设 A 有特征根 $\rho e^{i\theta} \neq \rho$，其中 $0 \leqslant \theta \leqslant 2\pi$. 在这批特征根中取这样的特征根，使得 θ 最小. 由定理 14.1.5，存在正实数构成之对角方阵 D，使得 $(D^{-1}AD)e = \rho e$. 由定理 14.1.4，

$$D^{-1}AD = e^{\sqrt{-1}\theta}(\exp\sqrt{-1}\Lambda_1)^{-1}D^{-1}AD(\exp\sqrt{-1}\Lambda_1),$$

其中 $\Lambda_1 = \text{diag}(\theta_{11}, \cdots, \theta_{1n}), 0 \leqslant \theta_{11}, \cdots, \theta_{1n} < 2\pi$. 由于 $\exp\sqrt{-1}\Lambda_1$ 为对角方阵，所以 $D(\exp\sqrt{-1}\Lambda_1) = (\exp\sqrt{-1}\Lambda_1)D$. 这证明了

$$A = e^{\sqrt{-1}\theta}(\exp\sqrt{-1}\Lambda_1)^{-1}A(\exp\sqrt{-1}\Lambda_1).$$

由 A 和 $e^{\sqrt{-1}\theta}A$ 相似，所以 A 有特征根

$$\rho,\ \rho e^{\sqrt{-1}\theta}, \rho e^{2\sqrt{-1}\theta}, \cdots.$$

但是 A 的特征根的个数有限，所以存在最小自然数 m 使得

$e^{\sqrt{-1}m\theta}=1$,即$e^{\sqrt{-1}\theta}$为$x^m=1$的本原根,即 $\theta=\frac{2\pi}{m}$. 于是记 $\varepsilon=e^{\frac{2\pi\sqrt{-1}}{m}}$,则$A$有特征根$\varepsilon\rho,\varepsilon^2\rho,\cdots,\varepsilon^m\rho$. 这时若$A$另有特征根 $\rho e^{\sqrt{-1}\varphi}$,其中$0<\varphi<2\pi$,又$\varphi\neq\frac{2j\pi\sqrt{-1}}{m}$,$j=1,2,\cdots,m$. 由$\theta$之选取可知$\frac{2\pi}{m}<\varphi$. 因此存在自然数$p$使得

$$\frac{2p\pi}{m}<\varphi<\frac{2(p+1)\pi}{m}.$$

此即$0<\varphi-\frac{2p\pi}{m}<\frac{2\pi}{m}$ 由于A和$e^{\frac{-2p\pi\sqrt{-1}}{m}}A$相似,所以从$A$有特征根$\rho e^{\sqrt{-1}\varphi}$可推出$A$有特征根$\rho e^{\sqrt{-1}\left(\varphi-\frac{2p\pi}{m}\right)}$,这和$\theta$之选取矛盾.所以形如$\rho e^{\sqrt{-1}\theta}$之特征根只有$\varepsilon\rho,\varepsilon^2\rho,\cdots,\varepsilon^m\rho$. 由定理14.1.5,$A$有单重根$\rho$. 由$A$和$e^{\sqrt{-1}\theta}A$相似,所以$\varepsilon\rho,\varepsilon^2\rho,\cdots,\varepsilon^m\rho$都是$A$的单重根.

由定理14.1.4,所以$D^{-1}AD$有属于特征根$\varepsilon j\rho$的特征向量$(\exp\sqrt{-1}\Lambda_j)^{-1}e$,其中$\Lambda_j=\mathrm{diag}(\theta_{j1},\cdots,\theta_{jn})$,$0\leqslant\theta_{j1},\theta_{j2},\cdots,\theta_{jn}<2\pi$. 今$D^{-1}AD(\exp\sqrt{-1}\Lambda_j)^{-1}e=\varepsilon j\rho(\exp\sqrt{-1}\Lambda_j)^{-1}e$,所以

$$A[(\exp\sqrt{-1}\Lambda_j)^{-1}De]=(\varepsilon j\rho)[\exp\sqrt{-1}\Lambda_j)^{-1}De].$$

这证明了A的属于特征根$\varepsilon j\rho$的特征向量为$(\exp\sqrt{-1}\Lambda_j)^{-1}\alpha$,其中$\alpha=(d_1,\cdots,d_n)'$,$j=1,2,\cdots,m$.

今$0\leqslant\theta_{11},\cdots,\theta_{1n}<2\pi$. 于是存在置换方阵$P$,使得

$$P^{-1}\Lambda_1P=\mathrm{diag}(\mu_1I^{(n_1)},\cdots,\mu_sI^{(n_s)}),$$

其中$\mu_1,\cdots,\mu_s$为区间$[0,2\pi]$中不同数,且无妨设$\mu_1=0$. 将A和Λ一样分块,则有

$$A=\begin{pmatrix}A_{11}\cdots A_{1s}\\ \vdots \qquad \vdots\\ A_{s1}\cdots A_{ss}\end{pmatrix}.$$

由 $A=e^{\frac{2\sqrt{-1}\pi}{m}}(\exp\sqrt{-1}\Lambda_1)^{-1}A\ (\exp\sqrt{-1}\Lambda_1)$,所以有

$$(e^{\sqrt{-1}\mu_i}-e^{\sqrt{-1}(\frac{2\pi}{m}+\mu_j)})A_{ij}=0,\quad i,j=1,2,\cdots,s.$$

取 $j=1$，于是由 $\mu_1=0$ 有 $(e^{\sqrt{-1}\mu_i}-e^{\frac{2\sqrt{-1}\pi}{m}})A_{i1}=0$. 由 A 不可分拆，所以 $A_{11},A_{21},\cdots,A_{s1}$ 中至少有一块不等于零，而 $A_{11}=0$. 显然，无妨设 $A_{21}\neq0$，即 $\mu_2=\dfrac{2\pi}{m}$，于是 $A_{31}=0,\cdots,A_{s1}=0$，取 $j=2$，于是由 $\mu_1=0,\mu_2=\dfrac{2\pi}{m}$ 有 $(e^{\sqrt{-1}\mu_i}-e^{\frac{4\sqrt{-1}\pi}{m}})A_{i2}=0$，所以 $A_{12}=0,A_{22}=0$. 由 A 不可分拆，所以 $A_{32},\cdots,A_{s2}$ 中至少有一块不为零. 显然，无妨设 $A_{32}\neq0$，即 $\mu_3=\dfrac{4\pi}{m}$. 这样依次讨论下去，便有

$$\mu_j=\frac{2(j-1)\pi}{m},\quad j=1,2,\cdots,s,s\leqslant m$$

且

$$A=\begin{pmatrix}0&0&\cdots&0&A_{1s}\\A_{21}&0&\cdots&0&A_{2s}\\0&A_{32}&\cdots&0&A_{3s}\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&A_{s,s-1}&A_{ss}\end{pmatrix},$$

其中 $A_{21}\neq0,\cdots,A_{s,s-1}\neq0$，而 $(e^{\sqrt{-1}\mu_j}-e^{\frac{2s\sqrt{-1}\pi}{m}})A_{is}=0,\quad 1\leqslant i\leqslant s$. 由 A 不可分折可知 $s=m$，且 $A_{2m}=0,\cdots,\ A_{mm}=0$. 所以证明了定理.

习题 14.1

1. 设 A 为 n 阶不可分拆非负方阵，试证 $(I+A)^{n-1}$ 仍为非负方阵，其中元素全由正实数构成.

2. 设 $A_1,A_2,\cdots$ 为不可分拆非负方阵序列. 记 $A_k=(a_{ij}^{(k)})$，$k=1,2,\cdots$. 假设序列 $\{a_{ij}^{(k)}\}$ 为单调递增序列，$1\leqslant i,j\leqslant n$. 则 A_k 的最大正特征根 $\rho^{(k)}$ 也单调递增.

3. 设不可分拆非负方阵A的最大正特征根$\rho<1$，B为非负方阵. 试证：存在唯一的正数b，使得$A+bB$为非负方阵，且$A+bB$的最大实特征根为1.

4. 设ρ为n阶不可分拆非负方阵A的最大正特征根，记

$$\rho_i(x)=\frac{e'_iAx}{e'_ix},\ \forall\quad e'_ix>0.$$

试证：

$$\min_{x_1>0,\cdots,x_n>0}\max_{1\leqslant j\leqslant n}\rho_j(x)=\rho=\max_{x_1>0,\cdots,x_n>0}\min_{1\leqslant j\leqslant n}\rho_j(x).$$

5. 设ρ为不可分拆非负方阵A的最大正特征根. 记α为单位向量，其坐标大于或等于零. 设$A\alpha\leqslant\rho\alpha$或$A\alpha\geqslant\rho\alpha$，试证：$A\alpha=\rho\alpha$.

6. 试严格证明：$\varepsilon_1\rho,\cdots,\varepsilon_m\rho$都是不可分拆非负方阵$A$的单重特征根.

§14.2 非负方阵

设A为n阶非负方阵，记$e=(1,\cdots,1)'$. 任取自然数k，记

$$A_k=A+\frac{1}{k}ee'=A+\frac{1}{k}\begin{pmatrix}1&\cdots&1\\ \vdots&&\vdots\\ 1&\cdots&1\end{pmatrix},$$

则A_k为不可分拆非负方阵. 所以A_k有最大正特征根ρ_k，对应的单位特征向量α_k由正实数构成. 今序列$\{\alpha_k\}$有收敛子序列，极限为α_0，即$\lim\limits_{k\to\infty}\alpha_{i_k}=\alpha_0$ 它仍为单位向量，元素由非负实数构成. 由$A_{i_k}\alpha_{i_k}=\rho_{i_k}\alpha_{i_k}$，双方取极限，便证明了序列$\{\rho_{i_k}\}$有极限$\rho_0$，且

$$A\alpha_0=\rho_0\alpha_0.$$

今$\rho_{i_k}>0$，所以$\rho_0\geqslant0$. 又由$\det(e_kI-A_k)=0$，取极限，有$\det(\rho_0I-A)=0$. 所以ρ_0为A的特征根. 另一方面，$\det(\lambda I-A_k)\to\det(\lambda I-A)$，所以当$k$充分大时，由多项式$\det(\lambda I-A_k)$和$\det(\lambda I-A)$这两个多项式决定的曲线的纵坐标充分接近. 由于多项式函数的特性，所以零点也充分接近. 这证明了A_k的最大实特征根ρ_k和A的最大实特征根充分接近. 由$\rho_{i_k}\to\rho_0$，这证明了ρ_0为A的最大正特征根，所以$\rho_0>0$. 由定理14.1.5可知ρ_kI-

A_k 的伴随方阵$A(\rho_k)$由正实数构成．由于 $\lim_{k\to\infty} A(\rho_{i_k})=A(\rho_0)$，所以 $A(\rho_0)$为非负方阵．

引理 14.2.1 设ρ_0为非负方阵A的最大实特征根，则当$\lambda\geqslant\rho_0$时 $A(\lambda)$ 和 $\frac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)$都是非负方阵．

证 令$(\lambda I-A)A(\lambda)=\det(\lambda I-A)\cdot I$．记 $A(\lambda)$的第 i 行，第 j 列元素为 $A_{ij}(\lambda)$．于是由

$$A(\lambda)+(\lambda I-A)\frac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)=\left(\sum_{j=1}^{n}A_{jj}(\lambda)\right)I,$$

有

$$(\lambda I-A)^{-1}A(\lambda)+\frac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)=\left(\sum_{j=1}^{n}A_{jj}(\lambda)\right)(\lambda I-A)^{-1}.$$

所以由$(\lambda I-A)^{-1}=(\det(\lambda I-A))^{-1}A(\lambda)$，有

$$\begin{aligned}&(\det(\lambda I-A))^{-1}A(\lambda)^2+\frac{d}{d\lambda}A(\lambda)\\&\quad=(\det(\lambda I-A))^{-1}(\mathrm{tr}A(\lambda))A(\lambda).\end{aligned}$$

即有

$$\begin{aligned}\frac{d}{d\lambda}A_{ij}(\lambda)&=(\det(\lambda I-A))^{-1}\left[\sum_{l=1}^{n}A_{ll}(\lambda)A_{ij}(\lambda)\right.\\&\quad\left.-\sum_{l=1}^{n}A_{il}(\lambda)A_{lj}(\lambda)\right]\\&=(\det(\lambda I-A))^{-1}\sum_{l=1}^{n}\det\begin{pmatrix}A_{ll}(\lambda) & A_{il}(\lambda)\\ A_{lj}(\lambda) & A_{ij}(\lambda)\end{pmatrix}\end{aligned}$$

记 $n-1$ 阶方阵 $\lambda I^{(n-1)}-A^{(l)}=\lambda I^{(n-1)}-A\begin{pmatrix}1\cdots(l-1)(l+1)\cdots n\\1\cdots(l-1)(l+1)\cdots n\end{pmatrix}$的伴随方阵的第 i 行第 j 列元素为 $A_{ij}^{(l)}(\lambda)$，则易证

$$\det\begin{pmatrix} A_{il}(\lambda) & A_{il}(\lambda) \\ A_{lj}(\lambda) & A_{ij}(\lambda) \end{pmatrix} = (\det(\lambda I - A))A_{ij}^{(l)}(\lambda).$$

这证明了

$$\frac{d}{d\lambda}A_{ij}(\lambda) = \sum_{l \neq i,j} A_{ij}^{(l)}(\lambda) \qquad i, j = 1, 2, \cdots, n.$$

下面用归纳法来证明引理. 对阶数作归纳法. 当 $n=1$, 则 $A(\lambda)\equiv 1$, $\frac{d}{d\lambda}A(\lambda)\equiv 0$. 所以 $A(\lambda)$ 及 $\frac{d}{d\lambda}A(\lambda)$ 都是非负方阵. 设对阶数 $\leqslant n-1$ 的非负方阵, 相应的伴随方阵及其导数在 $\lambda \geqslant \rho_0$ 时都是非负方阵. 记 $A^{(l)} = A\binom{1\ 2 \cdots (l-1)(l+1)\cdots n}{1\ 2 \cdots (l-1)(l+1)\cdots n}$ 的最大实特征根为 ρ_l.

今

$$\det(\lambda I - A) = (\lambda - a_{nn})A_{nn}(\lambda) - \sum_{j,k=1}^{n-1} a_{nj}a_{kn}A_{jk}^{(n)}(\lambda).$$

对 $A^{(n)}$ 用归纳法假设, 所以 $A_{jk}^{(n)}$ 是非负实数, 因此

$$\det(\rho_n I - A) = (\rho_n - a_{nn})A_{nn}(\rho_n) - \sum_{j,k=1}^{n-1} a_{nj}a_{kn}A_{jk}^{(n)}(\rho_n).$$

由 $A_{nn}(\rho_n) = \det(\rho_n I - A^{(n)}) = 0$, 所以 $\det(\rho_n I - A) \leqslant 0$. 但是 ρ_0 为 A 的最大实特征根, 即当 $\lambda_0 > \rho_0$, 有 $\det(\lambda_0 I - A) > 0$. 这证明了 $\rho_n \leqslant \rho_0$. 同理可证 $\rho_i \leqslant \rho_0$, $i = 1, 2, \cdots, n$. 所以

$$\max_{1\leqslant i\leqslant n} \rho_i \leqslant \rho_0.$$

另一方面, 由归纳法假设, 当 $\lambda \geqslant \rho_i$ 时有 $A^{(i)}(\lambda)$ 非负. 这证明了当 $\lambda \geqslant \rho_0$ 时有

$$\frac{d}{d\lambda}A_{ij}(\lambda) = \sum_{l\neq i,j} A_{ij}^{(l)}(\lambda) \geqslant 0.$$

所以 $\frac{d}{d\lambda}A(\lambda)$ 非负. 又由 $\frac{d}{d\lambda}A_{ij}(\lambda) \geqslant 0$, 所以 $A_{ij}(\lambda)$ 单调递增. 因此当 $\lambda \geqslant \rho_0$ 时有 $A_{ij}(\lambda) \geqslant A_{ij}(\rho_0) \geqslant 0$, $1 \leqslant i, j \leqslant n$. 这证明了

$A(\lambda)$当$\lambda \geqslant \rho_0$时为非负方阵. 引理证完.

推论 1 设ρ_0为不可分拆非负方阵A的最大实特征根,则当$\lambda \geqslant \rho_0$时$A(\lambda)$由正实数构成.

证 由定理14.1.5,$A(\rho_0)$中元素全为正实数,由$A_{ij}(\lambda) \geqslant A_{ij}(\rho_0) > 0$, $i, j=1,2,\cdots,n$,所以$A(\lambda)$中元素全为正实数. 证完.

推论 2 设ρ_0为非负方阵A的最大实特征根,则$\det(\lambda I-A) > 0, \forall \lambda > \rho_0$. 因此$(\lambda I-A)^{-1}$为非负方阵,$\forall \lambda > \rho_0$. 特别当$A$为不可分拆非负方阵时,$(\lambda I-A)^{-1}$由正实数构成,$\forall \lambda > \rho_0$.

证 今ρ_0为非负方阵A的最大实特征根,所以$\det(\lambda I-A)$在区间$(\rho_0, +\infty)$中大于零. 由引理14.2.1,$(\det(\lambda I-A))^{-1}A(\lambda)$为非负方阵,所以$(\lambda I-A)^{-1}$为非负方阵. 当$A$不可分拆时,由推论可知$(\lambda I-A)^{-1}$由正实数构成. 证完.

定理 14.2.1 设ρ_0为n阶非负方阵A的最大实特征根,则A不可分拆当且仅当$\rho_0 I-A$的伴随方阵$A(\rho_0)$的对角元素$A_{jj}(\rho_0) > 0, j=1,2,\cdots,n$.

证 设A不可分拆,由于推论1,$A(\rho_0)$的对角元素全是正实数. 反之,若$A(\rho_0)$的对角元素全是正实数,记$A^{(l)}$为A中划去第l行,第l列构成的$n-1$阶子方阵,则

$$A_{ll}(\rho_0)=\det(\rho_0 I-A^{(l)}) > 0.$$

所以$n-1$阶方阵$A^{(l)}$的最大实特征根$\rho_l < \rho_0$, $l=1,2,\cdots,n$. 如果n阶方阵A可分拆,则无妨设

$$A=\begin{pmatrix} B^{(s)} & C \\ 0 & D \end{pmatrix},$$

其中$0 < s < n$. 又无妨设B的最大实特征根为ρ_0,所以A中除去最后一行,最后一列所得之子矩阵$A^{(n)}=\begin{pmatrix} B & C_1 \\ 0 & D_1 \end{pmatrix}$的最大实特征

根 $\rho_n<\rho_0$. 但是 $\rho_0\leqslant\rho_n$. 这导出矛盾. 证完.

定义 设 A 为 n 阶不可分拆非负方阵,若 A 的最大实特征根 ρ_0 大于所有其他特征根之模,则 A 称为**本原方阵**, 否则称为**非本原方阵**.

引理 14.2.2 元素都是正数的方阵为本原方阵.

证 由定理 14.1.6,设 A 的元素都是正数. 记 ρ 为 A 的最大正特征根, 则模等于 ρ 的特征根为 $\varepsilon_1\rho_1,\cdots,\varepsilon_m\rho$, 其中

$$\varepsilon_j=\exp\frac{2\sqrt{-1}(j-1)\pi}{m},\ j=1,2,\cdots,m$$

且
$$A=e^{\frac{2\pi\sqrt{-1}}{m}}(\exp^{\sqrt{-1}\Lambda_1})^{-1}A(\exp\sqrt{-1}\Lambda_1),$$
其中 $\Lambda_1=\mathrm{diag}(\theta_{11},\cdots,\theta_{1n}),0\leqslant\theta_{11},\cdots,\theta_{1n}<2\pi$. 记 $A=(a_{ij})$,则有

$e^{\sqrt{-1}\theta_{1i}}a_{ij}=a_{ij}e^{\sqrt{-1}\left(\frac{2\pi}{m}+\theta_{1j}\right)}$, $i,j=1,2,\cdots,n$. 由 $a_{ij}>0,i,j=1,$

$2,\cdots,n$,所以 $e^{\sqrt{-1}\theta_{1i}}a_{ii}=a_{ii}e^{\sqrt{-1}\left(\frac{2\pi}{m}+\theta_{1i}\right)}$ 这证明了 $m=1$. 所以 A 的其他特征根 λ_0 必有 $|\lambda_0|<\rho$. 证完.

引理 14.2.3 设 n 阶方阵 A 有 s 个不同特征根 $\lambda_1,\lambda_2,\cdots,\lambda_s$,任取自然数 m,则有

$$A^m=\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}(\lambda I-A)^{-1}]_{\lambda=\lambda_j},$$

其中

$$m(\lambda)=(\lambda-\lambda_1)^{e_1}(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_s)^{e_s}$$

为 A 的极小多项式.

证 由于所求公式在相似下不改变, 所以我们可取 $A=\mathrm{diag}(J_1^{(f_1)},\cdots,J_t^{(f_t)})$ 为 Jordan 标准形,其中

$$J_k^{(f_k)}=\mu_kI^{(f_k)}+N^{(f_k)},k=1,2,\cdots,t,$$

又 $\mu_1,\cdots,\mu_t$ 中不同数为 $\lambda_1,\cdots,\lambda_s$. 由于

$$A^m=\mathrm{diag}(J_1^{\ m},\cdots,J_t^{\ m}),$$

$$\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}(\lambda I-A)^{-1}]_{\lambda=\lambda_j}$$

$$=\operatorname{diag}\left(\cdots,\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}(\lambda I-J_k)^{-1}]_{\lambda=\lambda_j},\cdots\right),$$

所以问题化为计算

$$\Omega_k=\sum_{j=1}^{s}\frac{1}{(e-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}((\lambda-\mu_k)I^{(f_k)}-N^{(f_k)})^{-1}]_{\lambda=\lambda_j}.$$

注意到当 $\mu_k\neq\lambda_j$，则由该项中有因式$(\lambda-\lambda_j)^{e_j}$，所以当 $\lambda=\lambda_j$ 时取值为零，因此存在指标 j，使得 $\mu_k=\lambda_j$，且

$$\Omega_k=\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}\lambda^m(\lambda-\lambda_j)^{e_j}[(\lambda-\lambda_j)I^{(f_k)}-N]^{-1}{}_{\lambda=\lambda_j}$$

$$=\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}\lambda^m\left(\sum_{l=0}^{f_k-1}(\lambda-\lambda_j)^{e_j-l-1}N^l\right)_{\lambda=\lambda_j}$$

$$=\frac{1}{(e_j-1)!}\sum_{l=0}^{f_k-1}\left(\frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j-l-1}]_{\lambda=\lambda_j}\right)N^l$$

$$=\frac{1}{(e_j-1)!}\sum_{l=0}^{f_k-1}\frac{m!(e_j-1)!}{(m-l)!l!}\lambda_j^{m-l}N^l$$

$$=\sum_{l=0}^{f_k-1}\binom{m}{l}\lambda_j^{m-l}N^l$$

$$=\sum_{l=0}^{m}\binom{m}{l}\lambda_j^{m-l}N^l=(\lambda_jI^{(f_k)}-N^{(f_k)})^m=J_k^m,k=1,2,\cdots,t.$$

引理证完.

定理 14.2.2 n 阶不可分拆非负方阵 A 为本原方阵当且仅当存在自然数 m，使得 A^m 的元素都是正数.

证 设 A 为本原方阵，$\lambda_1,\cdots,\lambda_s$ 为 A 的不同特征根，其中 $\lambda_1>|\lambda_2|\geqslant\cdots\geqslant|\lambda_s|$. 由引理 14.2.3，对任意自然数 m，有

$$A^m=\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\ \frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}(\lambda I-A)^{-1}]_{\lambda=\lambda_j}.$$

今 A 不可分拆，A 的最大正特征根 λ_1 为单重根，所以 $e_1=1$. 因此

$$A^m=[\lambda^m(\lambda-\lambda_1)(\lambda I-A)^{-1}]_{\lambda=\lambda_1}$$

$$+\sum_{j=2}^{s}\frac{1}{(e_{j-1})!}\ \frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}(\lambda I-A)^{-1}]_{\lambda=\lambda_j}.$$

所以

$$\lambda_1^{-m}A^m=[(\lambda-\lambda_1)(\lambda I-A)^{-1}]_{\lambda=\lambda_1}$$

$$+\sum_{j=2}^{s}\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}\left[\left(\frac{\lambda}{\lambda_1}\right)^m(\lambda-\lambda_j)^{e_j}(\lambda I-A)^{-1}\right]_{\lambda=\lambda_j}$$

取 $m\to\infty$，有

$$\lim_{m\to\infty}\frac{1}{\lambda_1^m}A^m=[(\lambda-\lambda_1)(\lambda I-A)^{-1}]_{\lambda=\lambda_1}$$

$$=[(\lambda-\lambda_1)\det(\lambda I-A)^{-1}A(\lambda)]_{\lambda=\lambda_1}.$$

今 $\det(\lambda I-A)=(\lambda-\lambda_1)(\lambda-\lambda_2)^{k_2}\cdots(\lambda_1-\lambda_s)^{k_s}$, 所以

$$\lim_{m\to\infty}\lambda_1^{-m}A^m=(\lambda_1-\lambda_2)^{-k_2}\cdots(\lambda_1-\lambda_s)^{-k_s}A(\lambda_1).$$

由于 $\det(\lambda I-A)$ 为实多项式，所以复根成对出现，而实根都小于 λ_1，所以证明了 $(\lambda_1-\lambda_2)^{-k_2}\cdots(\lambda-\lambda_s)^{-k_s}>0$. 由引理 14.2.1，$A(\lambda_1)$ 中元素都是正实数，所以证明了 $\lim\limits_{m\to\infty}\lambda_1^{-m}A^m$ 为 n 阶方阵，元素都是正实数，所以存在充分大的数 m，使得 $\lambda_1^{-m}A^m$ 中元素都是正实数，即 A^m 中元素都是正实数.

反之，若 A 为不可分拆非负方阵，且存在自然数 m，使得 A^m 中元素都是正实数. 由引理 14.2.2 所以 A^m 为本原方阵. 因此 A^m 之不同特征根 $\mu_1,\cdots,\mu_s$ 有 $\mu_1>|\mu_2|\geqslant\cdots\geqslant|\mu_s|$. 由于 A 之 n 个特征根为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则 A^m 之 n 个特征根为 $\lambda_1^m,\lambda_2^m,\cdots,\lambda_n^m$. 又 A 有一个最大正特征根 λ_1，所以 $\lambda_1=\sqrt[m]{\mu_1}$，因此 A 之任意特征根 $\lambda_j\neq\lambda_1$，则有 $|\lambda_j|=\sqrt[m]{|\mu_k|}<\sqrt[m]{\mu_1}=\lambda_1$. 由定义可知 A 为本原方

阵. 证完.

定理 14.2.3 设 ρ 为不可分拆非负方阵 A 的最大正特征根，则 A 为非本原的当且仅当存在置换方阵 P 及一个最大的自然数 m，其中 $m\geqslant 2$，且

$$P^{-1}A^mP=\operatorname{diag}(A_1,A_2,\cdots,A_m),$$

其中 $A_1,A_2,\cdots,A_m$ 都是本原方阵，且它们的最大正特征根都是 ρ_m，又 A 恰有 m 个模等于 ρ 的特征根

$$\varepsilon_1\rho,\varepsilon_2\rho,\cdots,\varepsilon_m\rho$$

其中 $$\varepsilon_j=\exp\frac{2(j-1)\pi\sqrt{-1}}{m},j=1,2,\cdots,m.$$

证 设 $P^{-1}A^mP=\operatorname{diag}(A_1,A_2,\cdots,A_m)$，其中 $A_1,\cdots,A_m$ 本原，且 $m\geqslant 2$. 所以 A^m 可分拆，因此对任意自然数 q，A^{mq} 可分拆. 所以不存在自然数 p，使得 A^p 中元素全是正数. 否则由 A^p 中元素全是正数，则 A^{mp} 中元素也全是正数，这和 A^{mp} 可分拆矛盾. 由定理 14.2.2，所以 A 非本原.

反之，若 A 为非本原的不可分拆非负方阵. 所以 A 有特征根 $\rho e^{\sqrt{-1}\theta}\neq\rho$，其中 ρ 为 A 的最大正特征根. 由定理 14.1.6，A 的模等于 ρ 的特征根为

$$\varepsilon_1\rho,\varepsilon_2\rho,\cdots,\varepsilon_m\rho,$$

其中 $m\geqslant 2$ 为自然数，$\varepsilon_j=\dfrac{2\sqrt{-1}^{(j-1)\pi}}{m}$，$j=1,2,\cdots,m$. 且存在置换方阵 P 使得

$$P^{-1}e^{\sqrt{-1}A_1}P=\operatorname{diag}(\varepsilon_1I^{(n_1)},\cdots,\varepsilon_mI^{(n_m)}),$$

$$P^{-1}AP=\begin{pmatrix}0&\cdots&0&A_{1m}\\A_{21}&\cdots&0&0\\&\ddots&\vdots&\vdots\\&&A_{m,m-1}&0\end{pmatrix}.$$

于是 $$P^{-1}A^mP=\mathrm{diag}(A_1,\cdots,A_m).$$

下面分别求 $A_1,A_2,\cdots,A_m$ 的最大正特征根. 今 A 不可分拆,所以 $P^{-1}AP$ 不可分拆. 因此存在向量 $\alpha=(d_1,\cdots,d_n)'$ 为属于特征根 ρ 的特征向量,其中 $d_1>0,\cdots,d_n>0$. 将 α 和 $P^{-1}AP$ 一样分块, 则有

$$\alpha'=(\alpha_1',\cdots,\alpha_m').$$

由 $P^{-1}AP\alpha=\rho\alpha$, 便有 $P^{-1}A^mP\alpha=\rho^m\alpha$, 所以有

$$A_j\alpha_j=\rho^m\alpha_j,\qquad j=1,2,\cdots,m$$

这证明了 ρ^m 为 A_j 之特征根, 又对 A 之任一特征根 μ_0, 则有 $|\mu_0|\leqslant\rho$. 所以 $P^{-1}A^mP$ 有特征根 μ_0^m, 而 $|\mu_0^m|\leqslant\rho^m$. 这证明了 ρ^m 为 A_j 之最大正特征根. 设 $|\mu_0|=\rho$, 即

$$\mu_0=e^{\frac{2(k-1)\pi\sqrt{-1}}{m}}\rho,$$

其中 $k\in\{1,2,\cdots,m\}$. 所以 $\mu_0^m=\rho^m$. 这证明了 A_j 中模为 ρ^m 之特征根 μ_0^m 只有 ρ^m, 因此 A_j 中其他特征根之模都小于 ρ^m. 由定义可知, 为了证 A_j 为本原方阵, 只要证 A_j 不可分拆就够了, 这里 $j=1,2,\cdots,m$.

今 A 不可分拆, 若 $A_1,\cdots,A_m$ 中至少有一个可分拆, 由定理 14.1.1, 于是存在置换方阵 P_1, 使得非负方阵

$$P_1^{-1}P^{-1}A^mPP_1=\begin{pmatrix}\begin{pmatrix}B_{11} & \cdots & B_{1p}\\ & \ddots & \vdots\\ & & B_{pp}\end{pmatrix} & \begin{pmatrix}B_{1,p+1} & \cdots B_{1q}\\ \vdots & \vdots\\ B_{p,p+1} & \cdots B_{pq}\end{pmatrix}\\ & \begin{pmatrix}B_{p+1,p+1} & & \\ & \ddots & \\ & & B_{qq}\end{pmatrix}\end{pmatrix},$$

其中 $B_{11},\cdots,B_{qq}$ 不可分拆, 所以 $q>m$, 记 $B=P_1^{-1}P^{-1}APP_1$, 则由 A 不可分拆可知 B 及 B' 不可分拆, 它们的最大正特征根都是 ρ,

于是存在属于特征根 ρ 之特征向量 α,β，它们都由正实数构成。由 $B\alpha=\rho\alpha$, $B'\beta=\rho\beta$. 将 α 和 β 按 B^m 一样分块方式分块为 $\alpha'=(\alpha_1',\cdots,\alpha_q')$, $\beta'=(\beta_1',\cdots,\beta_q')$.

由

$$B^m\alpha=\rho^m\alpha,\ \beta'B^m=\rho^m\beta'.$$

所以有

$$B_{qq}\alpha_q=\rho^m\alpha_q,\ \beta'_qB_{qq}+\sum_{j=1}^{p}\beta_j'B_{jq}=\rho^m\beta_q'.$$

所以

$$\rho^m\beta_q'\alpha_q=\beta_q'B_{qq}\alpha_q+\sum_{j=1}^{p}\beta_j'B_{jq}\alpha_q.$$

由 $B_{qq}\alpha_q=\rho^m\alpha_q$，即有

$$\sum_{j=1}^{p}\beta_j'B_{jq}\alpha_q=0.$$

由于 B_{jq} 为非负方阵，β_j,α_q 之坐标都大于零，所以有 $\beta_j'B_{jq}\alpha_q=0$, $1\leqslant j\leqslant p$，即有 $B_{1q}=0,\cdots,B_{pq}=0$. 依次讨论下去，于是同理可证

$$B^m=\mathrm{diag}\left(\begin{pmatrix}B_{11} & \cdots & B_{1p}\\ & \ddots & \vdots\\ & & B_{pp}\end{pmatrix},\ B_{p+1,p+1},\cdots,B_{qq}\right)$$

于是

$$\sum_{k=j}^{q}B_{jk}\alpha_k=\rho^m\alpha_j,\ \sum_{k=1}^{j}\beta_k'B_{kj}=\rho^m\beta_j',\quad 1\leqslant j\leqslant p.$$

比较 $j=1$，有

$$\rho^m\alpha_1=\sum_{k=1}^{p}B_{1k}\alpha_k,\quad \rho^m\beta_1'=\beta_1'B_{11}$$

所以

$$\rho^m \beta_1' \alpha_1 = \beta_1' B_{11} \alpha_1 = \sum_{k=1}^{p} \beta_1' B_{1k} \alpha_k.$$

这证明了 $\sum\limits_{k=2}^{p} \beta_1' B_{1k} \alpha_k = 0$. 同理证明了 $B_{1k}=0$, $k=2,\cdots,p$. 依次讨论下去,于是同理可证 $B_{jk}=0$, $1 \leqslant j < k \leqslant p$. 即有

$$B^m = \mathrm{diag}(B_{11}, \cdots, B_{qq}),$$

其中 $B_{11},\cdots,B_{qq}$ 都不可分拆, 且 $B_{jj}\alpha_j = \rho^m \alpha_j$, 即每个 B_{jj} 都以 ρ^m 为最大正特征根, 因此都是 B_{jj} 之单重特征根, 所以 B^m 有 q 重特征根 ρ^m.

另一方面, A 的模为 ρ 之特征根都是单重根, 且为 $\varepsilon_1\rho, \varepsilon_2\rho, \cdots, \varepsilon_m\rho$. 所以 B^m 之模为 ρ^m 之特征根为 $\varepsilon_1^m\rho_m, \cdots, \varepsilon_m^m\rho^m$, 即 B^m 有 m 重特征根 ρ^m. 这证明了 $q=m$, 和 $q>m$ 矛盾, 所以证明了 $A_1, \cdots, A_m$ 都不可分拆, 即本原. 定理证完.

习题 14.2

1. 设 ρ 为 n 阶非负方阵 A 的最大实特征根, 试证: $\lambda_0 > \rho$ 当且仅当 $\lambda_0 I - A$ 的所有主子式都大于零.

2. 试用归纳法证明: 设 A 为 n 阶实方阵, a 为实数, $aI+A$ 为非负方阵, 设 A 的顺序主子式 $A\binom{1\,2\,\cdots\,k}{1\,2\,\cdots\,k}$ 有 $(-1)^k A\binom{1\,2\,\cdots\,k}{1\,2\,\cdots\,k} > 0, k=1,2,\cdots,n$, 试证: A 的一切主子式 $A\binom{i_1\,i_2\,\cdots\,i_k}{i_1\,i_2\,\cdots\,i_k}$ 有 $(-1)^k A\binom{i_1\,i_2\,\cdots\,i_k}{i_1\,i_2\,\cdots\,i_k} > 0, 1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$, 由此推出第一题之条件可改为一切顺序主子式都大于零.

3. 设 ρ 为非负方阵 A 的最大实特征根. 设 A 的特征根 ρ_0 有 $|\rho_0| < \rho$. 试证: 属于 ρ_0 的初等因式也都是一次因式.

4. 记 ρ 为非负方阵 A 的最大实特征根. 设 ρ 为 $\det(\lambda I - A)$ 的单重根, 且 A 和 A' 的属于 ρ 的特征向量的坐标都是正数. 试证: A 为不可分拆非负方阵.

5. 设 A 为不可分拆非负方阵. 记

$$\det(\lambda I - A) = \lambda^n + a_1 \lambda^{n_1} + \cdots + a_s \lambda^{n_s},$$

其中 $n > n_1 > \cdots > n_s \geqslant 0, a_1 a_2 \cdots a_s \neq 0$. 记 m 为 $n-n_1, \cdots, n_{s-1}-n_s$ 的最大公

因式. 试证: A 恰有 m 个模等于 A 的最大正特征根的特征根，它们也都是单重特征根.

§ 14.3 随机方阵

定义 记 A 为 n 阶非负方阵, $e=(1,\cdots,1)'$. 如果 $Ae=e$, 即 A 有特征根 1, 对应特征向量为 e, 则 A 称为 n 阶随机方阵.

引理 14.3.1 不可分拆随机方阵的最大正特征根 $\rho=1$. 又若 A 为不可分拆非负方阵, 其中 ρ 为 A 的最大正特征根. 记 $\alpha=(d_1,\cdots,d_n)'$ 为 A 的属于 ρ 的特征向量, 其中 $d_1>0,\cdots,d_n>0$. 记

$$\Lambda=\mathrm{diag}(d_1,\cdots,d_n).$$

则 $\rho^{-1}\Lambda^{-1}A\Lambda=B$ 为随机方阵.

证 设 ρ 为不可分拆随机方阵 A 的最大正特征根. 由定理 14.1.5, 所以 $\min\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^{n}a_{ij}\leqslant\rho\leqslant\max\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^{n}a_{ij}$ 其中 $A=(a_{ij})$. 由定义, $Ae=e$, 即有 $\sum\limits_{j=1}^{n}a_{ij}=1$. 这证明了 $\rho=1$.

又设 A 为不可分拆非负方阵. $A\alpha=\rho\alpha$, 即 $\sum\limits_{j=1}^{n}a_{ij}d_j=\rho d_i$, $i=1,2,\cdots,n$. 于是

$$A\Lambda e=A\alpha=\rho\alpha=\rho\Lambda e,$$

即有 $(\Lambda^{-1}A\Lambda)e=\rho e$. 由于 $\Lambda^{-1}A\Lambda$ 仍为非负方阵. 由定义, $\rho^{-1}\Lambda^{-1}A\Lambda=B$ 为随机方阵. 证完.

显然, 设 A 为随机方阵, 对 A 用置换方阵作相似, 则仍得随机方阵.

定理 14.3.1 随机方阵 A 有标准形

$$B=\begin{pmatrix} \begin{pmatrix} A_{11} & \cdots & A_{1p} \\ & \ddots & \vdots \\ & & A_{pp} \end{pmatrix} & \begin{pmatrix} A_{1,p+1} & \cdots & A_{1m} \\ \vdots & & \vdots \\ A_{p,p+1} & \cdots & A_{pm} \end{pmatrix} \\ & \begin{pmatrix} A_{p+1,p+1} & & \\ & \ddots & \\ & & A_{mm} \end{pmatrix} \end{pmatrix},$$

其中对角块 $A_{11},\cdots,A_{mm}$ 都是不可分拆非负方阵，$A_{p+1,p+1},\cdots,A_{mm}$ 为不可分拆随机方阵. 设 A_{11}, $\cdots$, A_{pp} 之最大正特征根分别记作 $\rho_1,\cdots,\rho_p$, 则有 $\rho_1<1,\cdots,\rho_p<1$.

证 由于 B 为随机方阵. 将 $e=(1,\cdots,1,)'$ 和 B 一样分块，即记 $e'=(\alpha_1',\cdots,\alpha_m')$ 则由 $Be=e$ 有 $A_{jj}\alpha_j=\alpha_j$, $j=p+1$, $\cdots,m$. 由于 $\alpha_j=\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$, 这证明了 $A_{p+1,p+1},\cdots,A_{mm}$ 为不可分拆随机方阵. 下面取 $j\in\{1,2,\cdots,p\}$. 设 A_{jj} 为 t 阶不可分拆非负方阵，记作 $A_{jj}=(a_{kl})$. 由 $(A_{j,j+1},\cdots,A_{jm})\neq 0$, 以及

$$\sum_{u=j}^{m}A_{ju}\alpha_u=\alpha_j.$$

所以 $A_{jj}\alpha_j$ 的坐标 $\leqslant 1$. 且至少有一个坐标 <1. 所以 $\min\limits_{1\leqslant k\leqslant t}\sum\limits_{l=1}^{t}a_{kl}=a<1$, $\max\limits_{1\leqslant k\leqslant t}\sum\limits_{l=1}^{t}a_{kl}=b\leqslant 1$. 由定理 14.1.5, $a\leqslant\rho_j\leqslant b$, 其中 ρ_j 为 A_{jj} 的最大正特征根.

设 $\rho_j=1$, 于是 $b=1,a<\rho_j$. 但是定理 14.1.5证明了这时必须有 $\rho_j=a$. 这导出矛盾. 所以证明了 $\rho_j<1$. 定理证完.

定理 14.3.2 随机方阵 A 的属于特征根 1 的初等因式都是一次的.

证 由定理 14.3.2, 随机方阵 A 的特征根 1 都在对角块

$A_{p+1,p+1},\cdots,A_{mm}$中出现. 由于 $A_{p+1,p+1},\cdots,A_{mm}$ 都是不可分拆非负方阵，它们的最大正特征根都等于 1，由定理 14.1.5 可知 1 是单重特征根. 这证明了定理. 证完.

引理 14.3.2 n 阶复方阵 A 构成的序列$\{A^q\}$有极限 A_∞ 当且仅当 A 的特征根 λ_0 的模$|\lambda_0|\leqslant 1$. 当$|\lambda_0|=1$时必有 $\lambda_0=1$，且这时 $\lambda_0=1$ 的初等因式都是一次的.

证 显然无妨设A为 Jordan 标准形. 因此问题化为对 Jordan 块 $\lambda_0 I+N$ 来讨论. 今

$$(\lambda_0 I+N)^q=\sum_{l=0}^{q}\binom{q}{l}\lambda_0^{q-l}N^l,$$

其中 $N^{n-1}\neq 0, N^n=0$.

设$\{(\lambda_0 I+N)^q\}$ 之极限 A_∞ 存在，于是对角元素 λ_0^q 之极限存在. 因此$|\lambda_0|\leqslant 1$. 当$|\lambda_0|<1$ 时，λ_0^q 之极限为 0；当$|\lambda_0|=1$ 时，λ_0^q 之极限存在当且仅当 $\lambda_0=1$. 这时 $(\lambda_0 I+N)^q=(I+N)^q=\sum_{l=0}^{q}\binom{q}{l}N^l$. 当$N$之阶数 $n>1$, 则 N的系数 q 构成的序列之极限不存在. 这证明了 $n=1$，即 $\lambda_0 I+N$ 之初等因式$(\lambda-\lambda_0)^n$ 为一次的.

反之，若 $A=\lambda_0 I+N$ 的特征根 λ_0 有$|\lambda_0|\leqslant 1$. 当$|\lambda_0|=1$时有 $\lambda_0=1$, 这时$n=1$. 显然$\{(\lambda_0 I+N)^q\}$收敛. 证完.

定理 14.3.3 设随机方阵 A 为本原方阵, 则序列$\{A^q\}$的极限存在, 且极限为

$$\lim_{q\to\infty}A^q=\lim_{\lambda\to 1}(\lambda-1)(\lambda I-A)^{-1}=A_\infty.$$

证 由引理 14.3.2 及定理 14.2.3,

$$\lim_{q\to\infty}A^q=\lim_{q\to\infty}\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^q(\lambda-\lambda_j)^{e_j}(\lambda I-A)^{-1}]_{\lambda=\lambda_j},$$

其中 $\lambda_1,\cdots,\lambda_s$, 为不同复数, 又

$$m(\lambda)=(\lambda-\lambda_1)^{e_1}(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_s)^{e_s}$$

为极小多项式. 且 $|\lambda_j|\leqslant 1$, 当 $|\lambda_j|=1$ 时 $\lambda_j=1$, 且初等因式都是一次的.

今由 A 为随机方阵, 所以有 $Ae=e$, 因此存在特征根1. 所以无妨设 $\lambda_1=1$. 由 A 为本原方阵, 所以 $|\lambda_l|<1$, $2\leqslant l\leqslant s$. 今 A 相似于 Jordan 标准形 $J=\mathrm{diag}(J_1^{(f_1)},\cdots,J_t^{(f_t)})$, 即存在非异复方阵 Q, 使得 $Q^{-1}AQ=J$, 其中 J_k 之特征根为 $\lambda_1,\cdots,\lambda_s$ 之一. 设 $J_k^{(f_k)}$ 之特征根为 λ_l, 其中 $l\in\{2,\cdots,s\}$. 则由引理 14.2.3 之证明可知

$$\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^q(\lambda-\lambda_j)^{e_j}(\lambda I-J_k)^{-1}]_{\lambda=\lambda_l}$$

$$=\sum_{l=0}^{f_k-1}\binom{q}{l}\lambda_l^{q-l}N^l\xrightarrow{q\to\infty}0.$$

设 $J_k^{(f_k)}$ 之特征根为1. 由于初等因式为一次的, 所以 $f_k=1$, 从而 $J_k=1$. 故

$$\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^q(\lambda-\lambda_j)^{e_j}(\lambda I-J_k)^{-1}]_{\lambda=\lambda_j}$$

$$=\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^q(\lambda-\lambda_j)^{e_j}(\lambda-1)^{-1}]_{\lambda=\lambda_j}$$

$$=\frac{1}{(e_1-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}[\lambda^q(\lambda-1)^{e_j-1}]_{\lambda=1}=1.$$

所以

$$\lim_{q\to\infty}A^q=Q\Big(\lim_{q\to\infty}\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{d^{e_j-1}}{d\lambda^{e_j-1}}\lambda^q(\lambda-\lambda_j)^{e_j}$$

$$\times\mathrm{diag}[(\lambda I-J_1)^{-1},\cdots,(\lambda I-J_t)^{-1}]_{\lambda=\lambda_j}\Big)Q^{-1}.$$

为方便起见, 无妨设 $J_1,\cdots,J_r$ 的特征根为 1, $J_{r+1},\cdots,J_t$ 之特征根不是 1. 于是

$$\lim_{q\to\infty} A^q = Q\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q^{-1}.$$

另一方面,

$$\lim_{\lambda\to 1}(\lambda-1)(\lambda I-A)^{-1} = Q\lim_{\lambda\to 1}(\lambda-1)$$

$$\times \mathrm{diag}((\lambda I-J_1)^{-1},\cdots,(\lambda I-J_t)^{-1})Q^{-1}.$$

由于$J_1=1,\cdots,J_t=1$,所以

$$\lim_{\lambda\to 1}(\lambda-1)(\lambda I-A)^{-1} = Q\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix} Q^{-1}.$$

定理证完.

定理 14.3.4 设标准形

$$A=\begin{pmatrix} \begin{pmatrix} A_{11} & \cdots & A_{1p} \\ & \ddots & \vdots \\ & & A_{pp} \end{pmatrix} & \begin{pmatrix} A_{1,p+1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{p,p+1} & \cdots & A_{pm} \end{pmatrix} \\ & \begin{pmatrix} A_{p+1,p+1} & & \\ & \ddots & \\ & & A_{mm} \end{pmatrix} \end{pmatrix} = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}.$$

为随机方阵,即有 $Ae=e$. 设 $A_{p+1,p+1},\cdots,A_{mm}$ 为本原方阵. 则

$$\lim_{q\to\infty} A^q = \begin{pmatrix} 0 & (I-B)^{-1}C \\ 0 & I \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & Q_\infty \end{pmatrix}.$$

其中

$$Q_\infty = \mathrm{diag}(Q_{p+1},\cdots,Q_m).$$

这里 $Q_{p+1},\cdots,Q_m$ 分别定义为

$$\lim_{q\to\infty} A_{kk}^q = Q_k,\quad k=p+1,\cdots,m.$$

记 $\lambda I-A_{kk}$ 的伴随方阵为 $A_k(\lambda)$,则

$$Q_k = A_{kk}(1)\lim_{\lambda\to 1}(\lambda-1)\det(\lambda I-A_{kk})^{-1} = e\beta'_k,$$

其中 $e'\beta_k=1$,且 β_k 由常数构成,$k=p+1,\cdots,m$.

证 由定理 14.3.3 及定理假设 $A_{p+1,p+1},\cdots,A_{mm}$ 为随机本原方阵，所以

$$\lim_{q\to\infty}A_{kk}^q=\lim_{\lambda\to 1}(\lambda-1)(\lambda I-A_{kk})^{-1}=Q_k,\ k=p+1,\cdots,m.$$

由于 A 为随机方阵，且 $A_{p+1,p+1}$，$\cdots$，A_{mm} 为本原方阵. 由定理 14.3.1，所以 A 为本原方阵. 由定理 14.3.3，所以

$$\lim_{q\to\infty}A^q=\lim_{\lambda\to 1}(\lambda-1)(\lambda I-A)^{-1}.$$

今

$$(\lambda I-A)=\begin{pmatrix}\lambda I-B & -C\\ 0 & \lambda I-D\end{pmatrix},$$

且 B 之特征根之模小于 1，所以在 $\lambda=1$ 附近，$\lambda I-B$ 有逆. 因此

$$(\lambda I-A)^{-1}=\begin{pmatrix}\lambda I-B & -C\\ 0 & \lambda I-D\end{pmatrix}^{-1}$$

$$=\begin{pmatrix}(\lambda I-B)^{-1} & (\lambda I-B)^{-1}C(\lambda I-D)^{-1}\\ 0 & (\lambda I-D)^{-1}\end{pmatrix}.$$

由于 $\lim\limits_{\lambda\to 1}(\lambda-1)(\lambda I-B)^{-1}=0$，

$$\lim_{\lambda\to 1}(\lambda-1)(\lambda I-D)^{-1}=\mathrm{diag}(Q_{p+1},\cdots,Q_m)\hat{=}Q_\infty.$$

所以

$$\lim_{q\to\infty}A^q=\begin{pmatrix}0 & (I-B)^{-1}CQ_\infty\\ 0 & Q_\infty\end{pmatrix}=\begin{pmatrix}0 & (I-B)^{-1}C\\ 0 & I\end{pmatrix}\begin{pmatrix}0 & 0\\ 0 & Q_\infty\end{pmatrix}.$$

于是问题化为证明：设 A 为 n 阶随机本原方阵，则

$$\lim_{q\to\infty}A^q=A(1)\lim_{\lambda\to 1}(\lambda-1)\det(\lambda I-A)^{-1}=e\beta',$$

其中 $\beta\in R^n$，且 $\beta'e=1$，又 β 之坐标全为正实数.

事实上，由定理 14.3.3，$\lim\limits_{q\to\infty}A^q=\lim\limits_{\lambda\to 1}(\lambda-1)(\lambda I-A)^{-1}$. 由于

记 $\lambda I-A$ 之伴随方阵为 $A(\lambda)$，则有

$$(\lambda I-A)^{-1}=(\det(\lambda I-A))^{-1}A(\lambda).$$

所以记

$$A_\infty=\lim_{q\to\infty}A^q=\lim_{\lambda\to1}(\lambda-1)(\det(\lambda I-A))^{-1}A(\lambda)$$
$$=A(1)\lim_{\lambda\to1}(\lambda-1)\det(\lambda I-A)^{-1}.$$

另一方面，A 为本原方阵，所以它的特征多项式

$$\det(\lambda I-A)=(\lambda-1)(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_s)^{e_s}$$

其中 $|\lambda_j|<1$，$j=2,\cdots,s$，又 $1,\lambda_2,\cdots,\lambda_s$ 为 s 个不同的复数. 由于 $A^{q+1}=AA^q$，所以取 $q\to\infty$，有 $A_\infty=AA_\infty$. 因此由 $a=\lim\limits_{\lambda\to1}(\lambda-1)\det(\lambda I-A)^{-1}>0$ 及 $A(1)$ 之元素为正实数，所以 A_∞ 之元素都是正实数. 另一方面，记 A_∞ 之列向量分别为 $\beta_1,\cdots,\beta_n$. 所以有 $A_\infty=(\beta_1\cdots\beta_n)$. 因此有

$$A\beta_j=\beta_j,\quad j=1,2,\cdots,n.$$

因此 $\beta_1,\cdots,\beta_n$ 都是方阵 A 的属于特征根 1 之特征向量. 由于 A 为本原随机方阵，所以 A 的特征根 1 为单重根，且 e 为特征向量. 这证明了 $\beta_j=a_je,j=1,2,\cdots,n$. 其中 $a_1,\cdots,a_n$ 为正实数. 记 $\beta=(a_1\cdots a_n)'$，则

$$A_\infty=(\beta_1\cdots\beta_n)=(a_1e\cdots a_ne)=e\beta'.$$

由 $Ae=e$，于是 $A^qe=e$. 因此 $A_\infty e=e$，所以 $e\beta'e=e$. 这证明了 $\beta'e=1$. 定理证完.

习题 14.3

1. 试证：任一非负方阵必在广义初等方阵下相似于随机方阵.

2. 设 A 为不可分拆非负方阵，ρ 为 A 的最大正特征根. 设 A 的模为 ρ 之不同特征根恰有 m 个. 试证：序列 $\{A^q\}$ 中有 m 个收敛子序列，它们的极限分别为 $A_\infty,A_\infty A,\cdots,A_\infty A^{N-1}$. 试用这个结论来证明序列 $\left\{\frac{1}{m}\sum\limits_{q=0}^{m}A^q\right\}$ 的极限存在.

3. 设 A 及 A' 都是随机方阵，则 A 称为**双随机方阵**. 试证：(i) A 为双随

机方阵当且仅当 A' 为双随机方阵. (ii) 置换方阵为双随机方阵. (iii) 设 A 为双随机方阵，且 A 中不等于零的元素的个数不超过 $n+3$，则存在置换方阵 P_1 和 P_2，使得 $A=\lambda P_1+(1-\lambda)P_2$，其中 $\lambda\in[0,1]$.

第十五章　矩阵偶的标准形理论

§15.1　矩阵偶在相抵下的标准形

在这一章考虑一对矩阵同时化标准形的问题。为此，我们称一对同型矩阵 A,B 为**矩阵偶**

定义　两个 $n\times m$ 矩阵偶$\{A,B\}$和$\{C,D\}$称为**相抵的**，如果存在一个 n 阶非异方阵 P 和一个 m 阶非异方阵 Q，使得

$$C=PAQ,\quad D=PBQ.$$

即 λ 矩阵 $A+\lambda B$ 和$C+\lambda D$ 有

$$C+\lambda D=P(A+\lambda B)Q.$$

在 § 13.1　我们定义了 λ 矩阵的相抵，

定义　两个 $n\times m$ λ 矩阵 $A(\lambda)$和$B(\lambda)$称为**严格相抵的**，如果存在 n 阶常数非异方阵 P 和 m 阶常数非异方阵 Q，使得

$$B(\lambda)=PA(\lambda)Q.$$

所以有

引理 15.1.1　两个 $n\times m$ 矩阵偶(A,B)和(C,D)相抵当且仅当 λ 矩阵 $A+\lambda B$ 和$C+\lambda D$ 严格相抵.

这一节的目的，是给出矩阵偶在相抵下的标准形和全系不变量.

设 λ 矩阵 $A+\lambda B$ 的秩为 r. 于是 $r\leqslant\min(n,\ m)$. 我们在下面分情形讨论. (1) $n=m=r$, 即 $A+\lambda B$ 为非异 λ 方阵；(2)$n\geqslant r$; $m>r$; (3) $n>r$, $m\geqslant r$. 在第三种情形，考虑$m\times n\lambda$矩阵 $(A+\lambda B)'=A'+\lambda B'$，则又化为情形(2).

设 $n \geq r, m > r$.

这时考虑线性方程组

$$(A+\lambda B)x=0.$$

视 λ 为任意参数，则由齐次线性方程组理论可知它有非零解. 任取一个非零解

$$x=\begin{pmatrix}\dfrac{f_1(\lambda)}{g_1(\lambda)}\\ \vdots \\ \dfrac{f_m(\lambda)}{g_m(\lambda)}\end{pmatrix},$$

其中 $f_1(\lambda),\cdots,f_m(\lambda),g_1(\lambda),\cdots,g_m(\lambda)$ 为 λ 的多项式,且 $g_1(\lambda),\cdots,g_m(\lambda)$ 为非零多项式. 由于齐次线性方程组之解乘以非零多项式后仍为 $(A+\lambda B)x=0$ 之非零解. 所以经过通分，我们总可取非零解为

$$x=(f_1(\lambda),\cdots,f_m(\lambda))',$$

其中 $f_1(\lambda),\cdots,f_m(\lambda)$ 为 m 个不全为零的多项式,且我们约定 $f_1(\lambda),\cdots,f_m(\lambda)$ 互素. 记

$$d=\max(\deg f_1(\lambda),\cdots,\deg f_m(\lambda)),$$

则 d 称为这个解 x 的次数.

为了今后讨论方便，我们称上面由 m 个不全为零的多项式构成的解为齐次线性方程组 $(A+\lambda B)x=0$ 的解.

定义 设 $n\times m$ λ 矩阵 $A+\lambda B$ 之秩 $r<m$, 则齐次线性方程组 $(A+\lambda B)x=0$ 之所有非零解中存在一个次数最小的非零解. 这个次数 ε (为非负整数)称为 λ 矩阵 $A+\lambda B$ 的**最小列指标**. 设 $n\times m$ λ 矩阵 $A'+\lambda B'$ 的秩 $r<n$, 则它的最小列指标称为 $n\times m$ λ 矩阵 $A+\lambda B$ 的**最小行指标**.

引理 15.1.2 最小列(行)指标都在严格相抵下不改变.

证 设 $C+\lambda D=P(A+\lambda B)Q$, 其中 P,Q 为非异常数方阵. 则

齐次线性方程组$(A+\lambda B)x=0$可改写为$P^{-1}(C+\lambda D)Q^{-1}x=0$，即$(C+\lambda D)y=0$，其中$y=Q^{-1}x$. 这证明了$C+\lambda D$的最小列指标$d'\leqslant d$. 由于$x=Qy$，同理可证$d\leqslant d'$，所以证明了$d'=d$. 对最小行指标的证明相同. 引理证完.

引理 15.1.3 设ε为$n\times m$ λ矩阵$A+\lambda B$的最小列指标. 则存在$m\times 1$ λ矩阵

$$\alpha=\alpha_0+\lambda\alpha_1+\cdots+\lambda^{\varepsilon}\alpha_{\varepsilon},$$

其中$\alpha_0,\alpha_1\cdots,\alpha_{\varepsilon}\in C^m$，且$\alpha_{\varepsilon}\neq 0,\alpha_0\neq 0$，又$\alpha_0,\alpha_1,\cdots,\alpha_{\varepsilon}$线性无关，且$A\alpha_1,\cdots,A\alpha_{\varepsilon}\in C^n$也线性无关.

证 令$A+\lambda B$有非零解

$$\alpha=\begin{pmatrix}f_1(\lambda)\\ \vdots\\ f_m(\lambda)\end{pmatrix}=\alpha_0+\lambda\alpha_1+\cdots\lambda^{\varepsilon}\alpha_{\varepsilon},$$

其中$\alpha_{\varepsilon}\neq 0$. 由$f_1(\lambda),\cdots,f_m(\lambda)$互素可知$\alpha\neq 0$. 今$(A+\lambda B)\alpha=0$，所以有

$$A\alpha_0=0,A\alpha_1+B\alpha_0=0,\cdots,A\alpha_{\varepsilon}+B\alpha_{\varepsilon-1}=0,B\alpha_{\varepsilon}=0.$$

先来证当$\varepsilon>0$，有$A\alpha_1,A\alpha_2,\cdots,A\alpha_{\varepsilon}$线性无关. 今若$A\alpha_1=0$，于是$B\alpha_0=0$，因此$(A+\lambda B)\alpha_0=0$. 但是$\alpha_0\neq 0$. 这证明了最小列指标$\varepsilon=0$. 这和条件$\varepsilon>0$矛盾. 所以$A\alpha_1\neq 0$，即$A\alpha_1$线性无关. 设$A\alpha_1,\cdots,A\alpha_t$线性无关，且$t<\varepsilon$. 我们来证$A\alpha_1,\cdots,A\alpha_t,A\alpha_{t+1}$也线性无关. 设若不然，则$A\alpha_{t+1}$为$A\alpha_1\cdots,A\alpha_t$的线性组合，所以存在常数$a_0,\cdots,a_{t-1}$，使得

$$A\alpha_{t+1}+\alpha_{t-1}A\alpha_t+\cdots+a_0A\alpha_1=0.$$

由于$B\alpha_{j-1}=-A\alpha_j,j=1,2,\cdots,\varepsilon$. 代入，有

$$B(a_0\alpha_0+a_1\alpha_1+\cdots+a_{t-1}\alpha_{t-1}+\alpha_t)=0.$$

记

$$\beta_t=a_0\alpha_0+\alpha_1\alpha_1+\cdots+a_{t-1}\alpha_{t-1}+\alpha t,$$

则有 $B\beta_t=0$，又

$$A\beta_t=a_0A\alpha_0+a_1A\alpha_1+\cdots+a_{t-1}A\alpha_{t-1}+A\alpha_t$$
$$=-a_1B\alpha_0-a_2B\alpha_1-\cdots-a_{t-1}B\alpha_{t-2}-B\alpha_{t-1}=-B\beta_{t-1}$$

其中 $\qquad \beta_{t-1}=a_1\alpha_0+a_2\alpha_1+\cdots+a_{t-1}\alpha_{t-2}+\alpha_{t-1}$.

这样依次作下去，即令

$$\beta_j=a_{t-j}\alpha_0+\cdots+a_{t-1}\alpha_{j-1}+\alpha_j,\ j=0,1,\cdots,t$$

则由归纳法易证

$$B\beta_t=0,B\beta_{t-1}+A\beta_t=0,\cdots,B\beta_0+A\beta_1=0,A\beta_0=0.$$

作 λ 矩阵

$$\beta=\beta_0+\lambda\beta_1+\cdots+\lambda^t\beta_t,$$

则有 $(A+\lambda B)\beta=0$、注意到 $\beta_0=\alpha_0\neq0$，所以 β 为齐次线性方程组 $(A+\lambda B)x=0$ 之非零解，且 β 由 m 个不全为零的多项式构成，它的最大次数 $\leqslant t<\varepsilon$。这和最小列指标的定义矛盾. 所以我们证明了 $A\alpha_1,\cdots,A\alpha_\varepsilon$ 线性无关.

再证 $\alpha_0,\alpha_1,\cdots\alpha_\varepsilon$ 线性无关. 如果它们相关，则有一组不全为零的常数 $a_0,a_1,\cdots,a_\varepsilon$，使得 $\sum\limits_{l=0}^{\varepsilon}a_l\alpha_l=0$. 于是由 $A\alpha_0=0$，则有 $\sum\limits_{l=1}^{\varepsilon}a_lA\alpha_l=0$，上面已证 $A\alpha_1\cdots,A\alpha_\varepsilon$ 线性无关，所以证明了 $a_1=a_2=\cdots=a_\varepsilon=0$. 因此 $a_0\alpha_0=0$，但是 $\alpha_0\neq0$，所以 $a_0=0$. 这就导出矛盾，所以证明了 $\alpha_0,\alpha_1,\cdots,\alpha_\varepsilon$ 线性无关. 引理证完.

引进标准块

$$L_k=\begin{pmatrix}\lambda & -1 & & & \\ & \lambda & -1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & -1\end{pmatrix}^{(k,k+1)}=\lambda(I^{(k)}0^{(k,1)})-(0^{(k,1)}I^{(k)}).$$

则有

引理 15.1.4 设 ε 为 $n\times m$ λ 矩阵 $A+\lambda B$ 的最小列指标，且 $\varepsilon>0$. 则存在 n 阶非异常数方阵 P_1 及 m 阶非异常数方阵 Q_1，使得

$$P_1(A+\lambda B)Q_1=\begin{pmatrix} L_\varepsilon & 0 \\ 0 & A_1+\lambda B_1 \end{pmatrix},$$

其中 $A_1+\lambda B_1$ 为 $(n-\varepsilon)\times(n-\varepsilon-1)$ λ 矩阵，且若 $A_1+\lambda B_1$ 有最小列指标 τ，则有 $\tau\geqslant\varepsilon$.

证　由 ε 之定义，所以存在 $(A+\lambda B)x=0$ 之非零解

$$\alpha=\alpha_0+\lambda\alpha_1+\cdots+\lambda^\varepsilon\alpha_\varepsilon,$$

其中 $\alpha_0\neq0,\cdots,\alpha_\varepsilon\neq0$, 又 $\alpha_0,\alpha_1,\cdots,\alpha_\varepsilon\in C^m$.

今 $(A+\lambda B)\alpha=0$ 等价于

$A\alpha_0=0$, $A\alpha_1+B\alpha_0=0,\cdots,A\alpha_\varepsilon+B\alpha_{\varepsilon-1}=0, B\alpha_\varepsilon=0$.

记

$$\bar{\alpha}=\begin{pmatrix}\alpha_0\\ \alpha_1\\ \vdots\\ \alpha_\varepsilon\end{pmatrix}\in C^{(\varepsilon+1)m},$$

则有 $(A+\lambda B)\alpha=0$ 等价于

$$\begin{pmatrix} A & & & \\ B & A & & \\ & B & \ddots & \\ & & \ddots & A \\ & & & B \end{pmatrix}\bar{\alpha}=0$$

记 $(k+2)n\times(k+1)m$ 常数矩阵

$$M_k(A+\lambda B)=\begin{pmatrix} A & & & \\ B & A & & \\ & B & \ddots & \\ & & \ddots & A \\ & & & B \end{pmatrix}, k=0,1,\cdots,\varepsilon,$$

所以 $(A+\lambda B)\alpha=0$ 等价于 $M_\varepsilon(A+\lambda B)\ \alpha=0$. 所以 $M_\varepsilon(A+\lambda B)$ 的秩小于 $(\varepsilon+1)m$. 由最小列指标之定义可知 $M_{\varepsilon-1}(A+\lambda B)$ 之秩等于 εm. 所以 $M_k(A+\lambda B)$ 之秩为 $(k+1)m$, $k=0, 1, \cdots, \varepsilon-1$.

由此可见，我们可以这样来计算最小列指标 ε. 先求 $A+\lambda B$ 之秩 r，当 $r<m$. 再依次构造 $M_0(A+\lambda B)$, $M_1(A+\lambda B)$, $\cdots$,使它们的列都是满秩的. 直到第一个 $M_\varepsilon(A+\lambda B)$，它的列不满秩，那末 ε 就是 λ 矩阵 $A+\lambda B$ 之最小列指标.

现在开始证明引理. 由引理 15.1.1 $\alpha_0, \alpha_1, \cdots, \alpha_\varepsilon$ 线性无关. 所以存在 m 阶常数非异方阵 $Q_2=(\alpha_0\alpha_1\cdots\alpha_\varepsilon *)$ 又由 $A\alpha_1, A\alpha_2, \cdots, A\alpha_\varepsilon$ 线性无关,又可构造 n 阶常数非异方阵 $P_2^{-1}=-(A\alpha_1, \cdots, A\alpha_\varepsilon, *)$. 今

$$
\begin{aligned}
P_2(A+\lambda B)Q_2&=P_2((A+\lambda B)\alpha_0, \cdots, (A+\lambda B)\alpha_\varepsilon, *)\\
&=P_2(-\lambda A\alpha_1, A\alpha_1, -\lambda A\alpha_2, \cdots, A\alpha_{\varepsilon-1}-\lambda A\alpha_\varepsilon, A\alpha_\varepsilon, *)\\
&=P_2(-A\alpha_1, \cdots, -A\alpha_\varepsilon, *)\begin{pmatrix} L_\varepsilon & * \\ 0 & * \end{pmatrix}\\
&=\begin{pmatrix} L_\varepsilon & A_2+\lambda B_2 \\ 0 & A_1+\lambda B_1 \end{pmatrix}.
\end{aligned}
$$

现在来证明:若 $A_1+\lambda B_1$ 有最小列指标 τ，则 $\tau\geqslant\varepsilon$. 今无妨设

$$
A+\lambda B=\begin{pmatrix} L_\varepsilon & A_2+\lambda B_2 \\ 0 & A_1+\lambda B_1 \end{pmatrix},
$$

于是

$$
A=\begin{pmatrix} (0 \quad -I) & A_2 \\ 0 & A_1 \end{pmatrix}, \quad B=\begin{pmatrix} (I \quad 0) & B_2 \\ 0 & B_1 \end{pmatrix}.
$$

这里 A_1, B_1 为$(n-\varepsilon)\times(m-\varepsilon-1)$矩阵. 为了证 $\tau\geqslant\varepsilon$，我们只要证 $(\varepsilon+1)(n-\varepsilon)\times\varepsilon(m-\varepsilon-1)$矩阵. $M_{\varepsilon-1}(A_1+\lambda B_1)$ 之列为满

秩的．今已知 $A+\lambda B$ 的最小列指标为ε，即 $M_{\varepsilon-1}(A+\lambda B)$的列是满秩的．由于 A,B 之分块方式相同，由 $M_{\varepsilon-1}(A+\lambda B)$ 之定义，对 $M_{\varepsilon-1}(A+\lambda B)$ 用置换方阵作相抵，则可证 $M_{\varepsilon-1}(A+\lambda B)$相抵于

$$\begin{pmatrix} M_{\varepsilon-1}(L_\varepsilon) & M_{\varepsilon-1}(A_2+\lambda B_2) \\ 0 & M_{\varepsilon-1}(A_1+\lambda B_1) \end{pmatrix}$$

今 $\operatorname{rank}(M_{\varepsilon-1}(A+\lambda B))=\varepsilon m$，$\operatorname{rank}(M_{\varepsilon-1}(L_\varepsilon))=\varepsilon(\varepsilon+1)$，且 $M_{\varepsilon-1}(L_\varepsilon)$为$\varepsilon(\varepsilon+1)$阶方阵，所以它非异，所以

$$\operatorname{rank}(M_{\varepsilon-1}(A_1+\lambda B_1))=\varepsilon(m-\varepsilon-1).$$

由于 $M_{\varepsilon-1}(A_1+\lambda B_1)$为$(n-\varepsilon)(\varepsilon+1)\times\varepsilon(m-\varepsilon-1)$矩阵，所以，$M_{\varepsilon-1}(A_1+\lambda B_1)$之列是满秩的．这证明了 $\tau\geqslant\varepsilon$.

所以余下要证存在$\varepsilon\times(m-\varepsilon)$常数矩阵 X及 $(\varepsilon+1)\times(m-\varepsilon-1)$常数矩阵$Y$，使得

$$\begin{pmatrix} I & X \\ 0 & I \end{pmatrix}\begin{pmatrix} L_\varepsilon & A_2+\lambda B_2 \\ 0 & A_1+\lambda B_1 \end{pmatrix}\begin{pmatrix} I & Y \\ 0 & I \end{pmatrix}=\begin{pmatrix} L_\varepsilon & 0 \\ 0 & A_1+\lambda B_1 \end{pmatrix}.$$

此即求X,Y使得 $L_\varepsilon Y+(A_2+\lambda B_2)+X(A_1+\lambda B_1)=0$，或

$$(0\ -I)Y+A_2+XA_1=0,\ (I\ \ 0)Y+B_2+XB_1=0.$$

将 X,Y,A_2,B_2 按行分块为

$$X=\begin{pmatrix} u_1 \\ \vdots \\ u_\varepsilon \end{pmatrix},\quad Y=\begin{pmatrix} v_1 \\ \vdots \\ v_{\varepsilon+1} \end{pmatrix},\ A_2=\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_\varepsilon \end{pmatrix},\ B_2=\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_\varepsilon \end{pmatrix}.$$

则有

$$v_{j+1}=\alpha_j+u_jA_1,\quad v_j=-\beta_j-u_jB_1,\quad j=1,2,\cdots,\varepsilon.$$

所以有

$$v_j=\alpha_{j-1}+u_{j-1},\ \ A_1=-\beta_j-u_jB_1,\ j=2,\cdots,\varepsilon$$

$$v_{\varepsilon+1}=\alpha_\varepsilon+u_\varepsilon A_1,\ v_1=-\beta_1-u_1B_1.$$

即有

$$(u_1 \ \cdots \ u_\varepsilon)M_{\varepsilon-2}(A_1+\lambda B_1)=-(\alpha_1+\beta_2,\cdots,\alpha_{\varepsilon-1}+\beta_\varepsilon).$$

今$M_{\varepsilon-2}(A_1+\lambda B_1)$为$\varepsilon(n-\varepsilon)\times(\varepsilon-1)(m-\varepsilon-1)$矩阵，为$(\varepsilon-1)(m-\varepsilon-1)$. 视$(u_1\cdots u_\varepsilon)$为$\varepsilon(m-\varepsilon)$个独立未知数构成之向量. 所以这是$\varepsilon(m-\varepsilon)$个未知数，$(\varepsilon-1)(m-\varepsilon-1)$个方程构成之非齐次线性方程组. 而系数矩阵之秩和方程个数相同. 所以一定有解. 任取一组解，便求出了$u_1,\cdots,u_\varepsilon$，所以求出了X. 再由$v_j=-\beta_i-u_jB_1$, $j=1,2,\cdots,\varepsilon$, $v_{\varepsilon+1}=\alpha_\varepsilon+u_\varepsilon A_1$，便求出了$v_1,\cdots,v_{\varepsilon+1}$，所以求出了$Y$、至此证明了引理证完.

由此引理立即有

定理 15.1.1 给定$n\times m$矩阵偶A,B，则λ矩阵$A+\lambda B$严格相抵于

$$\begin{pmatrix} 0^{(\varepsilon_0,\eta_0)} & & & & & & & \\ & L\varepsilon_1 & & & & & & \\ & & \ddots & & & & & \\ & & & L\varepsilon_p & & & & \\ & & & & L'\eta_1 & & & \\ & & & & & \ddots & & \\ & & & & & & L'\eta_q & \\ & & & & & & & A_0+\lambda B_0 \end{pmatrix}.$$

其中$\varepsilon_0=0$, $0^{(0,\eta_0)}$表示它的前η_0列全为零，$\eta_0=0$, $0^{(\varepsilon_0,0)}$表示它的前ε_0行全为零. 又

$$\varepsilon_0,\eta_0\geqslant 0,\ 0<\varepsilon_1\leqslant\varepsilon_2\leqslant\cdots\leqslant\varepsilon_p,\ 0<\eta_1\leqslant\eta_2\leqslant\cdots\leqslant\eta_q$$

又$A_0+\lambda B_0$为n_0阶λ方阵，且$\det(A_0+\lambda B_0)\neq 0$.

证 由引理 15.1.4，依次作出$0^{(\varepsilon_0,\eta_0)}, L_{\varepsilon_1},\cdots,L_{\varepsilon_p}$. 于是余下之$\lambda$矩阵无最小列指标. 这时考虑它的转置矩阵，即考虑最小行指标. 于是最后所得之λ矩阵$A_0+\lambda B_0$无最小行指标，也无最小列指标. 所以$A_0+\lambda B_0$之秩r等于行数，也等于列数，即为非异λ方阵. 证完.

这样一来，求矩阵偶(A, B)在相抵下的标准形的问题化为求非异λ方阵$A_0+\lambda B_0$在严格相抵下的标准形.

定理 15.1.2 设$A_0+\lambda B_0$为n_0阶非异λ方阵. 则$A_0+\lambda B_0$严格相抵于标准形

$\Omega=\mathrm{diag}(I^{(\delta_1)}+\lambda N^{(\delta_1)},\cdots, I^{(\delta_s)}+\lambda N^{(\delta_s)}, (\lambda+\lambda_1)I^{(r_1)}+N^{(r_1)},\cdots,(\lambda+\lambda_u)I^{(r_u)}+N^{(r_u)})$，其中$0<\delta_1\leqslant\cdots\leqslant\delta_s$.

证 今$\det(A_0+\lambda B_0)\neq0$，即为$\lambda$之非零多项式，所以存在复数$\lambda$. 使得$\det(A_0+\lambda_0 B_0)\neq0$. 而

$A_0+\lambda B_0=A_0+\lambda_0B_0+(\lambda-\lambda_0)B_0=(A_0+\lambda_0B_0)[I+(\lambda-\lambda_0)(A_0+\lambda_0B_0)^{-1}B_0]$ 将常数方阵$(A_0+\lambda_0B_0)^{-1}B_0$相似于Jordan标准形$J$，这证明了$A_0+\lambda B_0$严格相抵于$\lambda J+I-\lambda_0J$. 将$J$中属于特征根零之Jordan块拼在一起，记作$J_0$，将$J$中不属于特征根零的Jordan块拼在一起，记作$J_1$. 于是$J=\mathrm{diag}(J_0,J_1)$，而

$$I-\lambda_0J+\lambda J=\begin{pmatrix}I-\lambda_0J_0+\lambda J_0 & 0\\ 0 & I-\lambda_0J_1+\lambda J_1\end{pmatrix}.$$

由$\det(I-\lambda_0J_0)=1$，$J_0(I-\lambda_0J_0)^{-1}$幂零，所以相似于Jordan标准形$\tilde{J}_0=\mathrm{diag}(N^{(\delta_1)},\cdots, N^{(\delta_s)})$. 由$\det J_1\neq0$，故$J_1^{-1}(I-\lambda_0J_1)$相似于Jordan标准形$\tilde{J}_1=\mathrm{diag}(\lambda_1I^{(r_1)}+N^{(r_1)},\cdots,\lambda_uI^{(r_u)}+N^{(r_u)})$. 所以$I-\lambda_0J+\lambda J$严格相抵于标准形$\Omega$. 定理证完.

于是，$n\times m$ λ矩阵$A+\lambda B$在严格相抵下的标准形为

$$\begin{pmatrix}0^{(\varepsilon_0,\eta_0)} &&&&&&& \\ & L\varepsilon_1 &&&&&& \\ && \ddots &&&&& \\ &&& L\varepsilon_p &&&& \\ &&&& L'\eta_1 &&& \\ &&&&& \ddots && \\ &&&&&& L'\eta_q & \\ &&&&&&& A_0+\lambda B_0\end{pmatrix},$$

其中

$A_0+\lambda B_0=\mathrm{diag}(I^{(\delta_1)}+\lambda N^{(\delta_1)},\cdots,I^{(\delta_s)}+\lambda N^{(\delta_s)},(\lambda+\lambda_1)I^{(r_1)}+N^{(r_1)}\cdots,(\lambda+\lambda_u)I^{(r_u)}+N^{(r_u)})$ 这时 $0\leqslant\varepsilon_0$，$0<\varepsilon_1\leqslant\cdots\leqslant\varepsilon_p$ 称为 $A+\lambda B$ 之行指标，$0\leqslant\eta_0$，$0<\eta_1\leqslant\cdots\leqslant\eta_q$ 称为 $A+\lambda B$ 之列指标. 又 $\lambda^{-\delta_1},\cdots,\lambda^{-\delta_s},(\lambda+\lambda_1)^{r_1},\cdots,(\lambda+\lambda_u)^{r_u}$，$0<\delta_1\leqslant\cdots\leqslant\delta_s$ 称为 $A+\lambda B$ 的全部初等因式.

定理 15.1.3 $n\times m$ λ 矩阵 $A+\lambda B$ 的全部行指标，列指标以及初等因式是它在严格相抵下的全系不变量. 且标准形中对角块 $(\lambda+\lambda_j)I^{(r_j)}+N^{(r_j)}$，$j=1,2,\cdots,u$ 若不计次序，则标准形唯一.

证 记

$$G_1=\begin{pmatrix}L\varepsilon_1 & & \\ & \ddots & \\ & & L\varepsilon_p\end{pmatrix},\quad G_2=\begin{pmatrix}L'\eta_1 & & \\ & \ddots & \\ & & L'\eta_q\end{pmatrix},\quad G_3=A_0+\lambda B_0$$

即 $A+\lambda B$ 在严格相抵下的标准形为

$$D=\begin{pmatrix}0^{(\varepsilon_0,\eta_0)} & & & \\ & G_1 & & \\ & & G_2 & \\ & & & G_3\end{pmatrix}=\begin{pmatrix}0^{(\varepsilon_0,\eta_0)} & 0\\ 0 & D_1\end{pmatrix}.$$

如果另有标准形

$$\tilde{D}=\begin{pmatrix}0^{(\tilde{\varepsilon}_0,\tilde{\eta}_0)} & & & \\ & \tilde{G}_1 & & \\ & & \tilde{G}_2 & \\ & & & \tilde{G}_3\end{pmatrix}=\begin{pmatrix}0^{(\tilde{\varepsilon}_0,\tilde{\eta}_0)} & 0\\ 0 & \tilde{D}_1\end{pmatrix}.$$

下面我们依次来证明(i) $\tilde{\varepsilon}_0=\varepsilon_0$，$\tilde{\eta}_0=\eta_0$，(ii) $\tilde{G}_1=G_1$，$\tilde{G}_2=G_2$；(iii) $\tilde{G}_3$ 由 G_3 重排对角块次序而得到.

(1) 证 $\tilde{\varepsilon}_0=\varepsilon_0$，$\tilde{\eta}_0=\eta_0$.

由于 D 及 $\tilde{D}$ 都是 $A+\lambda B$,在严格相抵下的标准形，即存在 n 阶常数非异方阵 P 及 m 阶常数非异方阵 Q，使得

$$\tilde{D}=PDQ.$$

而 $\dim\{\beta\in C^m|\tilde{D}\beta=0\}=\dim\{\beta\in C^m|PDQ\beta=0\}=\dim\{Q\beta\in C^m|D(Q\beta)=0\}=\dim\{\delta\in C^m|D\delta=0\}$，所以 $\tilde{\eta}_0=\eta_0$. 同理可证 $\tilde{\varepsilon}_0=\varepsilon_0$. 证完.

(2) 证 D_1 和 $\tilde{D}_1$ 严格相抵.

将 P 及 Q 分成四块，有

$$\begin{pmatrix} P_1^{(\varepsilon_0)} & P_2 \\ P_3 & P_4 \end{pmatrix}\begin{pmatrix} 0^{(\varepsilon_0,\eta_0)} & 0 \\ 0 & D_1 \end{pmatrix}\begin{pmatrix} Q_1^{(\eta_0)} & Q_2 \\ Q_3 & Q_4 \end{pmatrix}=\begin{pmatrix} 0^{(\varepsilon_0,\eta_0)} & 0 \\ 0 & \tilde{D}_1 \end{pmatrix}.$$

即 $\quad P_2D_1Q_3=0, P_2D_1Q_4=0, P_4D_1Q_3=0, P_4D_1Q_4=\tilde{D}_1.$

如果 $\det P_4=0$，则存在非零向量 $\alpha\in C^{n-\varepsilon_0}$ 使得 $\alpha'P_4=0$，于是 $\alpha'\tilde{D}_1=0$，即 $\tilde{D}_1'\alpha=0$. 由 ε_0, η_0 之定义可知 $\alpha=0$. 这导出矛盾，所以 $\det P_4\neq 0$. 同理可证 $\det Q_4\neq 0$，所以 D_1 和 $\tilde{D}_1$ 严格相抵. 证完.

(3) $\tilde{G}_1=G_1$，且 $\mathrm{diag}(G_2, G_3)$ 和 $\mathrm{diag}(\tilde{G}_2, \tilde{G}_3)$ 严格相抵.

由于最小列指标在严格相抵下不变. 所以 $\tilde{\varepsilon}_1=\varepsilon_1$. 设 $\varepsilon_1=\cdots=\varepsilon_{n_1}<\varepsilon_{n_1+1}$，$\tilde{\varepsilon}_1=\cdots=\tilde{\varepsilon}_{m_1}<\tilde{\varepsilon}_{m_1+1}$，无妨设 $m_1\leqslant n_1$. 所以可以记

$$D_1=\begin{pmatrix} L\varepsilon_1 & & & \\ & \ddots & & \\ & & L\varepsilon_{m1} & \\ & & & D_2 \end{pmatrix},\quad \tilde{D}_1=\begin{pmatrix} L\varepsilon_1 & & & \\ & \ddots & & \\ & & L\varepsilon_{m1} & \\ & & & \tilde{D}_2 \end{pmatrix}$$

则 D_2 和 $\tilde{D}_2$ 严格相抵，由归纳法，便证明了断言.

为此，将 P 及 Q^{-1} 分成四块，由 $PD_1=\tilde{D}_1Q^{-1}$ 有

$$\begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix}\begin{pmatrix} \begin{pmatrix} L\varepsilon_1 & & \\ & \ddots & \\ & & L\varepsilon_{m_1} \end{pmatrix} & 0 \\ 0 & D_2 \end{pmatrix}=\begin{pmatrix} \begin{pmatrix} L\varepsilon_1 & & \\ & \ddots & \\ & & L\varepsilon_{m_1} \end{pmatrix} & 0 \\ 0 & \tilde{D}_2 \end{pmatrix}\begin{pmatrix} Q_1^{(m_1\varepsilon_1)} & Q_2 \\ Q_3 & Q_4 \end{pmatrix},$$

所以有

$$P_1\begin{pmatrix}L\varepsilon_1 & & \\ & \ddots & \\ & & L\varepsilon_{m_1}\end{pmatrix}=\begin{pmatrix}L\varepsilon_1 & & \\ & \ddots & \\ & & L\varepsilon_{m_1}\end{pmatrix}Q_1,\quad P_2D_2=\begin{pmatrix}L\varepsilon_1 & & \\ & \ddots & \\ & & L\varepsilon_{m_1}\end{pmatrix}Q_2,$$

$$P_3\begin{pmatrix}L\varepsilon_1 & & \\ & \ddots & \\ & & L\varepsilon_{m_1}\end{pmatrix}=\tilde{D}_2Q_3,\quad P_4D_2=\tilde{D}_2Q_4.$$

作 ε_1+1 维复向量 α 使得 $\alpha'=(1,\lambda,\cdots\lambda^{\varepsilon_1})$，则 $L\varepsilon_1\alpha=0$. 取 m_1 个独立自变量 $x_1,\cdots,x_{m_1}$，则

$$\begin{pmatrix}L\varepsilon_1 & & \\ & \ddots & \\ & & L\varepsilon_{m_1}\end{pmatrix}\begin{pmatrix}x_1\alpha \\ \vdots \\ x_{m_1}\alpha\end{pmatrix}=0.$$

即有

$$\tilde{D}_2Q_3\begin{pmatrix}x_1\alpha \\ \vdots \\ x_{m_1}\alpha\end{pmatrix}=0.$$

但是$\tilde{D}_2$ 的最小列指标 $\varepsilon_{m_1+1}>\varepsilon_1$，而$Q_3\begin{pmatrix}x_1\alpha \\ \vdots \\ x_{m_1}\alpha\end{pmatrix}$之次数 小于等于 ε_1，这证明了它等于零. 注意到 λ 为独立参数，而 Q_3 为常数矩阵. 所以证明了$Q_3=0$. 于是$P_3\begin{pmatrix}L_{\varepsilon_1} & & \\ & \ddots & \\ & & L_{\varepsilon_{m_1}}\end{pmatrix}=0$. 因此证明了 $P_3=0$. 所以 $\det P_4\det Q_4\neq0$，即 D_2 和 $\tilde{D}_2$ 严格相抵. 证完.

(4) $\tilde{G}_2=G_2$，且 $\tilde{G}_3$ 和G_3 严格相抵.

今由(3)，$\mathrm{diag}(G_2, G_3)$ 和 $\mathrm{diag}(\tilde{G}_2, \tilde{G}_3)$ 严格相抵，所以 $\mathrm{diag}(G_2',$

G_3') 和 diag($\tilde{G}_2'$, $\tilde{G}_3'$)严格相抵. 由(3) 便证明了$\tilde{G}_2'=G_2$, 且 $\tilde{G}_3'$ 和 G_3' 严格相抵. 因此 $\tilde{G}_2=G_2$, 且 $\tilde{G}_3$ 和G_3严格相抵. 证完.

(5) 最后, 设非异 λ 方阵 $A_0+\lambda B_0$ 和 $C_0+\lambda D_0$ 严格相抵. 由 $\det(A_0+\lambda B_0)\det(C_0+\lambda D_0)\neq 0$, 所以存在复数 λ_0 使得$\det(A_0+\lambda_0 B_0)\det(C_0+\lambda_0 D_0)\neq 0$. 记 $\lambda-\lambda_0=\tau$. 于是

$$I+\tau B_0(A_0+\lambda_0 B_0)^{-1}=I+\tau E_0,\quad I+\tau D_0(C_0+\lambda_0 D_0)^{-1}=I+\tau F_0$$

严格相抵. 我们来证 E_0 和 F_0 相似. 事实上, 存在非异常数复方阵 P_0, Q_0 使得 $I+\tau F_0=P(I+\tau E_0)Q$. 于是 $PQ=I$, 即 $Q=P^{-1}$. 又 $F_0=PE_0Q=PE_0P^{-1}$, 即 F_0 和 E_0 相似. 由定理14.1.3 可知, $A_0+\lambda B_0$ 和 $C_0+\lambda D_0$ 分别由 E_0 及 F_0 之 Jordan 标准形派生出来. 由 E_0 和F_0 相似可知它们的标准形相同 (在特征根非零之 Jordan 块不计次序意义下)这证明了定理.

习题 5.1

1. 试求矩阵偶 $A=\alpha\beta', B=\alpha\gamma'$ 在相抵下的标准形, 其中 $\alpha\in C^n$, $\beta,\gamma\in C^m$. 又求 $A=I^{(n)}$, $B=\begin{pmatrix} & & 1\\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix}$在相抵下的标准形.

§15.2 复对称及复斜对称方阵偶在相合下的标准形

在第八章我们知道, 一个复方阵可唯一地分解为一个复对称方阵和一个复斜对称方阵的和,且对复方阵作相合,则分解关系式不变.所以考虑复方阵在相合下的标准形化为考虑方阵偶(S,K), $S'=S$, $K'=-K$ 在同一个非异复方阵 P 的相合下的标准形. 为此推广为

定义 两个 n 阶复方阵偶(A, B)和(C, D)称为**相合的**(复相

合的)，如果存在 n 阶非异复方阵 P，使得 $C=P'AP, D=P'BP$ ($C=\bar{P}'AP, D=\bar{P}'BP$). 当 A,B,C,D,P 都是实方阵时，则称为**实相合的.**

在这一节考虑 A,B 同时为对称方阵或同时为斜对称方阵时，它们在相合下的标准形.

定理 15.2.1 对称(斜对称)方阵偶(A,B)和(C,D)相合的必要且充分条件为 λ 对称（斜对称）方阵 $A+\lambda B$ 和 $C+\lambda D$ 严格相抵.

证 今(A,B)和(C,D)相合，即存在非异复方阵 P 使得 $P'AP=C, P'BP=D$，所以 $P'(A+\lambda B)P=C+\lambda D$，即 $A+\lambda B$ 和 $C+\lambda D$ 严格相抵.

反之，若存在非异复方阵 P,Q，使得 $P(A+\lambda B)Q=C+\lambda D$，双方取转置，即有 $Q'(A+\lambda B)P'=C+\lambda D$. 所以 $Q'(A+\lambda B)P'=P(A+\lambda B)Q$. 取 $P^{-1}Q'=R$，即有 $R(A+\lambda B)=(A+\lambda B)R'$. 由 § 13.3 可知，存在多项式 $p(\lambda)$，使得 $p(R)^2=R$. 今 $\det P\det Q\neq 0$，所以 $\det R\neq 0$，因此 $\det p(R)\neq 0$. 令 $T=Q'p(R)^{-1}$，则有

$$\begin{aligned}C+\lambda D&=P(A+\lambda B)Q=Q'R^{-1}(A+\lambda B)Q=Q'p(R)^{-2}(A+\lambda B)Q\\&=Q'p(R)^{-1}(A+\lambda B)p(R')^{-1}Q=T(A+\lambda B)T'.\end{aligned}$$

由 $\det T\neq 0$ 便证明了定理. 证完.

由这个定理可知，对称(斜对称)方阵偶(A,B)在相合下的全系不变量为(A,B)在相抵下的全系不变量. 利用这点，所以问题化为先求出对称(斜对称)方阵偶的全系不变量，再寻找一个最简单的对称(斜对称)方阵偶，它具有相同的全系不变量. 这样做就建立了对称(斜对称)方阵偶在相合下的标准形理论.

引理 15.2.1 设 λ 方阵 $A+\lambda B$ 对称(斜对称)，则它在严格相抵下的标准形为

$\operatorname{diag}(0^{(r_0)}, L_{\varepsilon_1},\cdots,L_{\varepsilon_p},L'_{\varepsilon_1},\cdots,L'_{\varepsilon_p},(I+\lambda N)^{(\delta_1)},\cdots,$

$(I+\lambda N)^{(\delta_s)}, (\lambda+\lambda_1)I^{(r_1)}+N, \cdots, (\lambda+\lambda_u)I^{(r_u)}+N)$.

证 **由定理 15.1.1**，$A+\lambda B$ 及 $(A+\lambda B)'=\pm(A+\lambda B)$ 分别严格相抵于

$$\begin{pmatrix} 0^{(\varepsilon_0,\eta_0)} & & & & & & & \\ & L_{\varepsilon_1} & & & & & & \\ & & \ddots & & & & & \\ & & & L_{\varepsilon_p} & & & & \\ & & & & L'_{\eta_1} & & & \\ & & & & & \ddots & & \\ & & & & & & L'_{\eta_q} & \\ & & & & & & & A_0+\lambda B_0 \end{pmatrix},$$

$$\begin{pmatrix} 0^{(\eta_0,\varepsilon_0)} & & & & & & & \\ & L'_{\varepsilon_1} & & & & & & \\ & & \ddots & & & & & \\ & & & L'_{\varepsilon_1} & & & & \\ & & & & L_{\eta_1} & & & \\ & & & & & \ddots & & \\ & & & & & & L_{\eta_q} & \\ & & & & & & & A'_0+\lambda B'_0 \end{pmatrix}.$$

这证明了 $\varepsilon_0=\eta_0, \varepsilon_1=\eta_1, \cdots, \varepsilon_p=\eta_p, q=p$. 引理证完.

定理 15.2.2 对称 λ 方阵 $A+\lambda B$ 在相合下的标准形为

$$\Omega=\operatorname{diag}\left(0^{(\varepsilon_0)}, \begin{pmatrix} 0 & L_{\varepsilon_1} \\ L'_{\varepsilon_1}, & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & L_{\varepsilon p} \\ L'_{\varepsilon p} & 0 \end{pmatrix}, (I+\lambda N)S^{(\delta_1)}, \cdots,\right.$$

$$\left.(I+\lambda N)S^{(\delta_s)},\ ((\lambda+\lambda_1)I+N)S^{(r_1)}, \cdots, ((\lambda+\lambda_u)I+N)S^{(r_u)}\right),$$

其中

$$S=\begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix}.$$

证 由引理 15.2.1,所以 $A+\lambda B$ 严格相抵于

$$\operatorname{diag}\left(0^{(\varepsilon_0)},\ \begin{pmatrix}0 & L_{\varepsilon_1}\\ L'_{\varepsilon_1} & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & L_{\varepsilon_p}\\ L'_{\varepsilon_p} & 0\end{pmatrix},\ A_0+\lambda B_0\right)$$

其中 $A_0+\lambda B_0$ 仍对称. 由定理 15.1.1, 所以 $A+\lambda B$ 严格相抵于对称 λ 方阵 Ω. 由定理 15.2.1, $A+\lambda B$ 和 Ω 相合. 证完.

引理 15.2.2 给定Jordan 标准形 $J=\operatorname{diag}(J_1^{(e_1)},\cdots,J_t^{(e_t)})$, 其中 $J_k^{(e_k)}=\lambda_k I^{(e_k)}+N^{(e_k)}$, $k=1,2,\cdots,t$. 如果存在非异斜对称方阵 K, 使得 $KJ'=JK$, 则 J 的初等因式成对出现.

证 将 K 和 J 一样分块, 记作

$$K=\begin{pmatrix}K_{11} & \cdots & K_{1t}\\ \vdots & & \vdots\\ K_{t1} & \cdots & K_{tt}\end{pmatrix},$$

于是有 $J_jK_{jk}=K_{jk}J'_k$, $j,k=1,2,\cdots,t$. 今若

$$\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1\\ & & & \lambda\end{pmatrix}\begin{pmatrix}b_{11} & \cdots & b_{1q}\\ \vdots & & \vdots\\ b_{p1} & \cdots & b_{pq}\end{pmatrix}=\begin{pmatrix}b_{11} & \cdots & b_{1q}\\ \vdots & & \vdots\\ b_{p1} & \cdots & b_{pq}\end{pmatrix}\begin{pmatrix}\mu & & & \\ 1 & \mu & & \\ & \ddots & \ddots & \\ & & 1 & \mu\end{pmatrix},$$

即有

$$\lambda b_{i-1,j}+b_{ij}=\mu b_{i-1,j}+b_{i-1,j+1},\quad i=2,\cdots,p,\ j=1,\cdots,\ q-1.$$

$$\lambda b_{iq}+b_{i+1,q}=\mu b_{iq},\quad i=1,\cdots,p-1,$$

$$\lambda b_{pj}=\mu b_{pj}+b_{p,j+1},\quad j=1,2,\cdots,q-1$$

$$(\lambda-\mu)b_{pq}=0.$$

当 $\lambda\neq\mu$, 则证明了 $b_{kl}=0$, $1\leqslant k\leqslant p, 1\leqslant l\leqslant q$; 当 $\lambda=\mu$, 则有

$$b_{p2},\cdots,b_{pq}=0,\quad b_{2q}=\cdots=b_{pq}=0,\quad b_{ij}=b_{i-1,j+1},$$

$$i=2,\cdots,p,\ j=1,\cdots,q-1$$

所以证明了当 $\lambda_j\neq\lambda_k$, 则 $K_{jk}=0$; 当 $\lambda_j=\lambda_k$, 则有

$$K_{jk}=\begin{pmatrix}\begin{pmatrix}b_{jk}^{(1)} & \cdots & b_{jk}^{(\min(e_j,e_k))}\\ \vdots & \ddots & \\ b_{jk}^{(\min(e_j,e_k))} & & \end{pmatrix} & 0\\ 0 & 0\end{pmatrix}.$$

所以将特征根相同之 Jordan 块放在一起，从而无妨设 J 有相同的特征根，即有 $\lambda_1=\cdots=\lambda_t=\lambda_0$. 这时无妨设 $e_1\geqslant\cdots\geqslant e_t>0$. 特别 K_{jj} 为对称方阵. 由 $K=-K'$，所以 $K_{jj}=0$. 今若 $e_1>e_2$. 则 $(K_{11}\cdots K_{1t})$之第 e_1 行全为零，这和K非异矛盾. 所以 $e_1=e_2$. 设 $e_1=\cdots=e_r>e_{r+1}\geqslant\cdots\geqslant e_t>0$. 考虑$(K_{11}\cdots K_{1r}K_{1,r+1}\cdots K_{1t})$. 由 $K_{11}=0$，$(K_{1,r+1}\cdots K_{1t})$之第 e_1 行为零，所以 $(K_{12}\cdots K_{1r})$ 之第 e_1 行不等于零. 因此存在 K_{1j} 非异. 对 Jordan 块整块作初等变换，所以无妨设 K_{12} 非异. 因此 $2e_1$ 阶斜对称方阵$\begin{pmatrix}0 & K_{12}\\ -K'_{12} & 0\end{pmatrix}$非异，

记

$$K=\begin{pmatrix}\begin{pmatrix}0 & K_{12}\\ -K'_{12} & 0\end{pmatrix} & U\\ -U' & K_1\end{pmatrix},\quad J_0=\begin{pmatrix}N^{(e_1)} & 0\\ 0 & N^{(e_1)}\end{pmatrix},$$

$$\tilde{J}=\begin{pmatrix}N^{(e_3)} & & \\ & \ddots & \\ & & N^{(e_t)}\end{pmatrix}.$$

则由 $JK=KJ'$ 有 $\mathrm{diag}(N^{(e_1)},\cdots,N^{(e_t)})K=K\mathrm{diag}(N^{(e_1)},\cdots,N^{(e_t)})'$，所以有

$$J_0\begin{pmatrix}0 & K_{12}\\ -K'_{12} & 0\end{pmatrix}=\begin{pmatrix}0 & K_{12}\\ -K'_{12} & 0\end{pmatrix}J'_0,\quad J_0U=U\tilde{J}',\ \tilde{J}K_1=K_1\tilde{J}'.$$

而

$$\det K=\det\begin{pmatrix}0 & K_{12}\\ -K'_{12} & 0\end{pmatrix}\det\widetilde{K},$$

其中

$$\widetilde{K}=K_1+U'\begin{pmatrix}0 & K_{12}\\ -K'_{12} & 0\end{pmatrix}^{-1}U.$$

所以 $\widetilde{K}$ 为非异斜对称方阵. 又

$$\widetilde{J}\widetilde{K}=\widetilde{J}K_1+\widetilde{J}U'\begin{pmatrix}0 & K_{12}\\ -K'_{12} & 0\end{pmatrix}^{-1}$$

$$U=K_1\widetilde{J}'+U'\begin{pmatrix}0 & K_{12}\\ -K'_{12} & 0\end{pmatrix}^{-1}U\widetilde{J}'=\widetilde{K}\widetilde{J}.$$

对 Jordan 标准形 $\widetilde{J}$ 和非异斜对称方阵同样讨论,便证明了 $e_3=e_4$. 所以由归纳法便证明了引理. 证完.

定理 15.2.3 斜对称 λ 方阵 $A+\lambda B$ 在相合下的标准形为

$$\Omega=\operatorname{diag}\left(0^{(e_0)},\ \begin{pmatrix}0 & L_{e_1}\\ -L'_{e_1} & 0\end{pmatrix},\right.$$

$$\cdots,\begin{pmatrix}0 & L_{e_p}\\ -L_{e_p} & 0\end{pmatrix},\ \begin{pmatrix}0 & (I+\lambda N)S^{(\delta_1)}\\ -(I+\lambda NS)^{(\delta_1)} & 0\end{pmatrix},$$

$$\cdots,\ \begin{pmatrix}0 & (I+\lambda N)S^{(\delta_s)}\\ -(I+\lambda N)S^{(\delta_s)} & 0\end{pmatrix},$$

$$\begin{pmatrix}0 & ((\lambda+\lambda_1)I+N)S^{(r_1)}\\ -((\lambda+\lambda_1)I+N)S^{(r_1)} & 0\end{pmatrix},$$

$$\left.\cdots,\ \begin{pmatrix}0 & ((\lambda+\lambda_u)I+N)S^{(r_u)}\\ -((\lambda+\lambda_u)I+N)S^{(r_u)} & 0\end{pmatrix}\right).$$

证 显然, $A+\lambda B$ 严格相抵于

$$\operatorname{diag}\left(0^{(e_0)},\ \begin{pmatrix}0 & L_{e_1}\\ -L'_{e_1} & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & L_{e_p}\\ -L'_{e_p} & 0\end{pmatrix},\ A_0+\lambda B_0\right),$$

其中 $A_0+\lambda B_0$ 为非异斜对称 λ 方阵. 取复数 λ_0, $\det(A_0+\lambda_0 B_0)\neq 0$, 则 $A_0+\lambda B_0$ 严格相抵于 $I+(\lambda-\lambda_0)J$, 其中 J 为 Jordan 标准形. 即存在非异常数复方阵 P_0, Q_0, 使得 $A_0+\lambda B_0=P_0(I+(\lambda-\lambda_0)J)Q_0$. 由 $(A_0+\lambda B_0)'=-(A_0+\lambda B_0)$, 即有 $P_0(I+(\lambda-\lambda_0)J)Q_0$

$=-Q_0'(I+(\lambda-\lambda_0)J')P_0'$，即 $P_0Q_0=-Q_0'P_0'$，$P_0JQ_0=-Q_0'J'P_0'$. 所以 $J[Q_0(P_0^{-1})']=-P_0^{-1}Q_0'J'=[Q_0(P_0^{-1})']J'$. 今 $K=Q_0(P_0^{-1})'$ 斜对称，即有 $JK=KJ'$. 由引理 15.2.2，所以 J 之初等因式成对出现. 这证明了 $A+\lambda B$ 和 Ω 严格相抵. 由定理 15.2.1 及 $A+\lambda B$，Ω 同为斜对称 λ 方阵，所以 $A+\lambda B$ 和 Ω 相合. 定理证完.

上面的定理，是在复数范围内讨论. 在实的情形，由于一个实方阵 A 不一定能表成实多项式 $p(A)$ 之平方. 所以定理 15.2.1 不能形式推广. 因此对实的对称方阵及斜对称方阵偶，在实相合下的标准形比较复杂. 同样，对 Hermite 方阵偶在复相合下的标准形，也因为复方阵 A 虽可开方为 $p(A)$，但是 $\overline{p(A)}$ 一般不等于 $p(\bar{A})$. 因此定理15.2.1 仍不能直接推广. 这些标准形理论是经典结果. 但已超出本书范围. 下面仅叙述结果.

定理 15.2.4 实对称 λ 方阵 $A+\lambda B$(或 Hermite λ 方阵 $A+\lambda B$)的行指标和列指标相等. 又初等因式组为

$$\lambda^{-\delta_1},\cdots,\lambda^{-\delta_s},(\lambda+\lambda_1)^{r_1},\cdots,(\lambda+\lambda_u)^{r_u},(\lambda+\mu_1)^{\rho_1},\cdots,$$
$$(\lambda+\mu_t)^{\rho_t},(\lambda+\bar{\mu}_1)^{\rho_1},\cdots,(\lambda+\bar{\mu}_t)^{\rho_t}.$$

其中 $\lambda_1,\cdots,\lambda_u$ 实，$\mu_1,\cdots,\mu_t$ 复，且虚部不等于零.

定理 15.2.5 实对称方阵偶(A, B)决定之 λ 方阵 $A+\lambda B$ 在实相合下的标准形为实对称 λ 方阵. 它是准对角形. 对角块由下面五种标准形块构成:

$$0^{(e_0)},\ \begin{pmatrix}0 & L_\eta\\ L_\eta' & 0\end{pmatrix},\quad \varepsilon(I+\lambda N)S^{(\delta)},\quad \varepsilon((\lambda+\lambda_0)I+N)S^{(\xi)},$$

$$\varepsilon\rho\begin{pmatrix}(\lambda I+M\cos\theta)S^{(r)} & -(M\sin\theta)S^{(r)}\\ -(M\sin\theta)S^{(r)} & -(\lambda I+M\cos\theta)S^{(r)}\end{pmatrix},$$

其中 λ_0 为实特征根，$\rho e^{\sqrt{-1}\theta}$ 为复特征根，$\rho>0$，$0\leqslant\theta<2\pi$. 又 ε 取 ± 1 两值之一，称为标签，

$$M=\begin{pmatrix}1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1\end{pmatrix}=I+N.$$

于是实对称方阵偶的行指标组以及附以标签的初等因式组为它在实相合下的全系不变量.

定理 15.2.6 Hermite 方阵偶(A, B)决定之λ方阵$A+\lambda B$在复相合下的标准形为 Hermite λ方阵，它是准对角形. 对角块由下面五种标准块构成:

$$0^{(\epsilon_0)},\ \begin{pmatrix}0 & L_\eta\\ L'_\eta & 0\end{pmatrix},\ \varepsilon(I+\lambda N)S^{(\delta)},\ \varepsilon((\lambda+\lambda_0)I+N)S^{(\xi)},$$

$$\varepsilon\begin{pmatrix}0 & ((\lambda+\mu_0)I+N)S^{(\tau)}\\ ((\lambda+\overline{\mu}_0)I+N)S^{(\tau)} & 0\end{pmatrix},$$

$$S=\begin{pmatrix} & & & 1\\ & & \cdot^{\cdot^{\cdot}} & \\ & 1 & & \\ 1 & & & \end{pmatrix},$$

其中λ_0为实特征根，μ_0为复特征根，ε取± 1两值之一，称为标签. 于是 Hermite 方阵偶的行指标组以及附以标签的初等因式组为它在复相合下的全系不变量.

习题 15.2

1. 试证: n阶复方阵A和B复相合当且仅当复方阵偶$\left(\frac{1}{2}(A+A'),\ \frac{1}{2}(A-A')\right)$和$\left(\frac{1}{2}(B+B'),\ \frac{1}{2}(B-B')\right)$相抵.

附录一 数学归纳法

数学中的任何一个分支，总是先给出若干定义，即给出研究的对象，再界定若干公理. 而主要结果则构成这个分支的一套理论. 大致规律为: 最初是一切从定义出发，得出一系列基本结果，引出了一系列最基本的处理技巧. 再从定义和基本结果，导出一系列新的结果. 从数理逻辑的角度已经证明: 从一个确定的公理系统出发，可以导出无限多个结果. 但是作为主要理论，一般还是由有限多个重要结论构成. 这里附带强调一点，我们并没有论及为什么要建立一个分支？一个分支发展的推动力是什么？为什么有的分支成为主流？为什么有的分支逐渐失去它的黄金时代.

回到原来的问题. 数学的结果，总是从已知的结果或者条件出发，通过形式逻辑的推理，从而获得许多新的结果. 常用的推理方法(即数学证明的方法)可以分为演绎和归纳两大类.

演绎法有两种不同的形式:

(*a*) 直接证明.

它的格式可以写为: 因为…，所以…，所以…，…，等等. 因此证明了所需要的结果.

(*b*) 反证法.

它的格式可以写为: 设所需要的结果不成立，则…，于是…，所以…，这就导出矛盾. 因此证明了所需要的结果.

概括地说，直接证法是从已给的条件出发，通过各种已知的性质，一步一步地推出所要的断言. 逻辑格式是简单的，但是具体的推导，是需要各种数学技巧. 而反证法则是从已给的条件和假设所要证的断言不成立这两前提出发，通过各种已知的性质，尽可能地去寻找矛盾. 所以从原则上来讲，每一个可以用直接证法证明的

定理，一定可以用反证法来证明．反之亦然．然而，有时用直接证法极易导得的结果，用反证法就显得很复杂，反之亦然．所以，在具体证明一个性质时，应该选择一种最有力的证明方法．值得注意的是，对有些比较复杂的定理，在推导过程中，往往将直接证法和反证法交替使用，以便尽快地获得要证的结论．

具体证明一个性质时，有时也采用穷举法．这个方法的原则是：从已知条件出发，可以得到很多种可能的情形．将它们全部列出来，分别进行讨论，从而可以导出所有结论．

例 试证：实系数二次多项式 x^2+ax+b 不可能有一个实根 α 和一个复根 $\beta+\sqrt{-1}\gamma$，其中 $\gamma\neq0$．

直接证法．用公式解可知 x^2+ax+b 有两个根 $\frac{-a\pm\sqrt{a^2-4b}}{2}$ 于是判别式 $a^2-4b=\Delta$ 出现三种可能（这时用了穷举法）$\Delta>0$；$\Delta=0$；$\Delta<0$．在 $\Delta>0$ 或 $\Delta=0$ 时 $\frac{-a\pm\sqrt{\Delta}}{2}$ 为实数，所以两个根都是实数；在 $\Delta<0$ 时 $\frac{-a\pm\sqrt{\Delta}}{2}$ 都是复数，且虚部不等于零所以两根都不是实数．证完．

反证法．如果断言不成立，则 x^2+ax+b 有根 α 及 $\beta+\sqrt{-1}\gamma$，其中 $\gamma\neq0,\beta,\alpha$ 都是实数．所以

$$\begin{aligned}x^2+ax+b&=(x-\alpha)(x-\beta-\sqrt{-1}\gamma)\\&=x^2-(\alpha+\beta+\sqrt{-1}\gamma)x+\alpha(\beta+\sqrt{-1}\gamma).\end{aligned}$$

由于 $a=-(\alpha+\beta+\sqrt{-1}\gamma)$ 是实数，这就和 $\gamma\neq0$ 矛盾．故证完．

下面再介绍归纳法．归纳法有两种；有限数学归纳法和超限数学归纳法．后者可以在实变函数论中找到；前者有两种不同的形式，它们可以分别叙述为：

第一数学归纳法．如果性质 $P(n)$ 在 $n=1$ 时成立，而且在假设了性质 $P(n)$ 在 $n=k$ 时成立，经过逻辑推理可以证明性质 $P(n)$

在 $n=k+1$ 时也成立，那末性质 $P(n)$ 对一切自然数 n 都成立。

第二数学归纳法. 如果性质 $P(n)$ 在 $n=1$ 时成立，而且在假设了性质 $P(n)$ 在 $n\leqslant k$ 时都成立，经过逻辑推理可以证明性质 $P(n)$ 在 $n=k+1$ 时也成立，那末性质 $P(n)$ 对一切自然数 n 都成立.

下面分别举一例来说明它们的用法.

例 1 试证：对一切自然数 n，性质 $P(n)$：

$$1+2+\cdots+n=\frac{n(n+1)}{2}$$

成立.

证 取 $n=1$，显然性质 $P(1)$：$1=\frac{1\times 2}{2}$ 成立. 假设性质 $P(k)$ 成立，即有

$$1+2+\cdots+k=\frac{k(k+1)}{2}$$

由于

$$1+2+\cdots+k+(k+1)=\frac{k(k+1)}{2}+(k+1)=\frac{(k+1)(k+2)}{2},$$

所以性质 $P(k+1)$ 也成立. 由第一数学归纳法原理，所以性质 $P(n)$ 对一切自然数成立.

例 2 设 $a_1=3$，$a_2=7$，又

$$a_n=3a_{n-1}-2a_{n-2},\quad n=3,4,\cdots$$

试证：对一切自然数 n，总有

$$a_n=2^{n+1}-1.$$

证 今 $a_1=2^{1+1}-1$，$a_2=2^{2+1}-1$. 假设当 $n=1,2,\cdots,k$ 时有 $a_n=2^{n+1}-1$. 于是

$$a_{k-1}=2^k-1,\quad a_k=2^{k+1}-1.$$

因此

$$a_{k+1}=3a_k-2a_{k-1}=3(2^{k+1}-1)-2(2^k-1)=2^{k+2}-1.$$

由第二数学归纳法原理便证明了断言.

数学归纳法是不是合理的方法呢？事实上，它们不是公理，不是人为规定的. 从自然数的公理系统出发，我们可以证明数学归纳法是它们的推论.

定理 1 在自然数集合$\{1,2,\cdots\}$中任取非空子集合$\mathfrak{S}$，则在$\mathfrak{S}$中必有一个最小自然数.

证 今$\mathfrak{S}$非空，所以在$\mathfrak{S}$中存在一个数n_0，由于小于n_0的自然数构成集合$\mathfrak{S}_0=\{1,2\cdots,n_0\}$. 由$n_0\in\mathfrak{S}_0\cap\mathfrak{S}$，所以

$$\mathfrak{S}_0\cap\mathfrak{S}=\{1\leqslant n_{i_1}<n_{i_2}\cdots<n_{i_s}=n_0\}.$$

因此n_{i_1}为$\mathfrak{S}$中最小自然数. 证完.

定理 2 第一数学归纳法是合理的推理方法.

证 用反证法. 假设第一数学归纳法不是合理的推理方法，因此存在一种性质$P(n)$，它有$P(1)$成立，且若假设性质$P(k)$成立，则可推出$P(k+1)$成立. 但是性质$P(1)$，$P(2)$，$\cdots$ 不是对所有自然数都成立. 因此使性质$P(n)$不成立的自然数构成自然数集$\{1,2,\cdots\}$之非空子集$\mathfrak{S}$. 由定理 1，其中存在最小自然数m. 于是$P(m)$不成立，但是$P(1),P(2),\cdots,P(m-1)$都成立. 由归纳法假设可知由$P(m-1)$成立便推出$P(m)$成立. 这导出矛盾. 所以证明了定理. 证完.

注意 这个定理的实质为：由于从$P(1)$成立可推出$P(2)$成立，$\cdots$，由$P(k)$成立可推出$P(k+1)$成立，$\cdots$，所以性质$P(1)$，$P(2),P(3),\cdots$都成立.

下面再证明第二数学归纳法只是第一数学归纳法的另一种表达形式，即有

定理 3 第一数学归纳法和第二数学归纳法互相等价.

证 假设性质$P(1)$成立. 于是问题化为证明："由性质$P(k)$

成立可以推出性质 $P(k+1)$成立”的必要且充分条件为“由性质 $P(1),P(2),\cdots,P(k)$都成立可以推出性质 $P(k+1)$成立”.

显然前一断言可以推出后一断言. 下面证明后一断言可以推出前一断言. 用反证法. 今若由性质 $P(1),\cdots,P(k)$ 成立便可推出 $P(k+1)$成立. 但是由 $P(k)$成立不一定能推出 $P(k+1)$成立. 取出所有自然数 k,使得由 $P(k)$成立推不出 $P(k+1)$成立. 它们全体构成自然数集的非空子集 $\mathfrak{S}$. 由定理 1,在 $\mathfrak{S}$ 中存在最小自然数 m. 因此从 $P(m)$成立推不出$P(m+1)$成立,但是由$P(1)$成立可推出$P(2)$成立,由 $P(2)$成立可推出$P(3)$成立,$\cdots$, 由 $P(m-1)$成立可推出 $P(m)$成立. 总之,$P(1),P(2),\cdots,P(m)$ 都成立. 由此便知 $P(m+1)$成立. 这和m的选取矛盾. 证完.

由此可知,第二数学归纳法也是合理的推理方法.

数学归纳法可以扩充到双指标,甚至多指标的情形. 我们只给出常用的双指标的情形.

双重数学归纳法: 考虑依赖于独立的自然数 m 及 n 的性质 $P(m,n)$. 若性质 $P(1,1)$成立,且由性质 $P(k,1)$ 成立 (或由性质 $P(1,1),P(2,1),\cdots,P(k,1)$都成立)可以推出性质 $P(k+1,1)$ 成立. 因此性质 $P(m,1)$ 对一切自然数都成立. 再假设性质 $P(m,k)$ 对一切自然数m都成立 (或假设性质 $P(m,1)$, $P(m,2),\cdots,P(m,k)$对一切自然数 m 都成立),则可推出性质 $P(m,k+1)$对一切自然数m 都成立. 那末性质 $P(m,n)$ 对一切自然数m和n都成立.

定理 4 双重数学归纳法是合理的推理方法.

证 由第一(第二)数学归纳法可知性质 $P(m,1)$对所有自然数m成立. 如果由性质 $P(m,k)$对一切自然数 m都成立可以推出性质 $P(m,k+1)$对一切自然数m都成立. 但是存在自然数 m_0,n_0 使得性质 $P(m_0,n_0)$不成立. 在集合

$$\{(m,n)\mid m,n=1,2,3,\cdots\}$$

中取出使得 $P(m,n)$ 不成立的所有自然数对 (m,n)，它们全体构成非空子集合 $\mathfrak{S}$. 在子集合 $\mathfrak{S}$ 中的自然数对中只取出后一个自然数，它们构成自然数集 $\{1,2,\cdots\}$ 之非空子集合 $\mathfrak{S}_1$. 由定理 1 可知在 $\mathfrak{S}_1$ 中存在最小自然数 n_1，显然 $n_1>1$. 由 n_1 之选取可知性质 $P(m,\ n_1-1)$ 对一切自然数 m 都成立. 所以推出性质 $P(m,n_1)$ 对一切自然数 m 都成立. 这和 n_1 选取矛盾. 所以证明了定理. 证完.

习题

1. 试直接证明第二数学归纳法为合理的推理方法.
2. 对于哪些自然数 n，不等式

$$2^n>n^3$$

成立？

3. 试证下列代数恒等式：

(1) $(x_1+\cdots+x_m)^n=\sum\limits_{i_1+\cdots+i_m=n}\dfrac{n!}{i_1!\,i_2!\cdots i_n!}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$,

(2) $\sum\limits_{k=0}^{n}(-1)^k\dbinom{n}{k}(n-k+1)^n=n!$.

附录 2 等价关系

我们知道，有理数可以用许多不同形式的分数 $\dfrac{p}{q}$ 来表示，而两个分数 $\dfrac{p_1}{q_1}$ 和 $\dfrac{p_2}{q_2}$ 表示同一有理数的必要且充分条件是 $p_1q_2=p_2q_1$，这时记作 $\dfrac{p_1}{q_1}=\dfrac{p_2}{q_2}$，这就引出了分数相等的概念. 所以相等的分数表示了同一个有理数. 将所有相等的分数放在一起，构成了实数的一个子集合，于是所有分数按照相等便分成了许多子集合. 每个子集合中的分数表示了同一个有理数，而不同集合中的分数表

示了不同的有理数. 所以这给出了子集合和有理数之间的一个自然的一一对应关系.

再例如固定向量的相等，矩阵的相等，矩阵的相抵，方阵的相似，数的相等，多项式的相等等等，都具有分数相等所导出的共性. 将这许许多多"相等"概念抽象化，便得到

定义 给定一个集合 $\mathfrak{S}$. 在集合中的元素之间定义了一种关系，记作$\cong$. 如果这种关系适合下面三条公理:

(1) **反身性**: 对集合 $\mathfrak{S}$ 中任一元素 a，有$a\cong a$;

(2) **对称性**: 对集合 $\mathfrak{S}$ 中任两元素 a,b，如果 $a\cong b$，则有 $b\cong a$;

(3) **传递性**: 对集合 $\mathfrak{S}$ 中任三元素 a, b, c，如果 $a\cong b$，$b\cong c$，则有 $a\cong c$.

则$\cong$称为**等价关系**. 当 $a\cong b$ 时，称元素 a 和 b **等价**.

由定义可知，前面提到的关系全是等价关系. 下面举出一些不是等价关系的关系的例子:

例 1 实数的顺序关系 $a\leqslant b$ 不是等价关系，事实上，它适合公理(1)，(3)，但不适合公理(2). 这个例子也证明了公理(2)和(1)，(3)独立.

例 2 实数的不等关系 $a<b$ 不是等价关系，事实上，它适合公理(3)，但不适合公理(1)，(2). 这个例子也证明了公理(3)和公理(1)，(2)独立.

例 3 实数 a 和 b 若有 $a+b=0$，则称 a 和 b 有一种关系. 这个关系也不是等价关系. 事实上，它适合公理(2)，但不适合公理(1)和(3). 这个例子也证明了公理(2)和公理(1)，(3)独立.

例 4 正实数 a 和 b 如果有 $ab<a+b^2$，则它定义了 a 和 b 间的一个关系. 这个关系也不是等价关系. 因为公理(1)适合，但不适合公理(2)，(3). 事实上，若 $ab<a+b^2$，不一定推得出 $ba<$

$b+a^2$. 例如取 $a=2,b=4$. 又若 $ab<a+b^2,bc<b+c^2$,不一定推得出 $ac<a+c^2$. 例如取 $a=5,b=1,c=2$.

等价关系是一种很重要的关系. 由等价关系出发，可以得到下面一些重要的性质.

定义 在集合 $\mathfrak{S}$ 中引进等价关系"$\cong$". 在集合 $\mathfrak{S}$ 中任取元素 a，把和 a 等价的元素放在一起构成一个子集合称为元素 a 的**等价类**,记作 $\mathfrak{S}_a$. a 称为等价类 $\mathfrak{S}_a$ 的**代表元素**.

引理 1 以 a 为代表元素的等价类有性质:

(1) $\mathfrak{S}_a$ 中任两元素等价;

(2) 不在 $\mathfrak{S}_a$ 中任一元素和 $\mathfrak{S}_a$ 中任一元素决不等价.

证 (1) 在 $\mathfrak{S}_a$ 中任取两元素 b,c. 由等价类 $\mathfrak{S}_a$ 的定义可知 $b\cong a,c\cong a$. 由对称性,有 $b\cong a,a\cong c$,由传递性,有 $b\cong c$. 所以性质(1)成立.

(2) 在 $\mathfrak{S}_a$ 中任取元素 b, 在 $\mathfrak{S}_a$ 外任取元素 c. 由等价类 $\mathfrak{S}_a$ 的定义可知 $b\cong a,c\ncong a$. 用反证法来证明性质(2). 设若不然，即存在 $b\in\mathfrak{S}_a,c\in\mathfrak{S}\setminus\mathfrak{S}_a$, 使得 $b\cong c$. 由对称性有 $b\cong a$, $c\cong b$. 由传递性,有 $c\cong a$. 这和 $c\ncong a$ 矛盾. 证完.

由这个引理可知，集合 $\mathfrak{S}$ 中元素 a 的等价类 $\mathfrak{S}_a$ 实际上是 $\mathfrak{S}_a$ 中任一元素的等价类,即有

引理 2 $\mathfrak{S}_a=\mathfrak{S}_b$ 当且仅当 $a\cong b$. 所以等价类 $\mathfrak{S}_a$ 的构成和元素 a 的选取无关,而是由等价关系完全确定的.

定理 1 在集合 $\mathfrak{S}$ 中给定等价关系$\cong$后，$\mathfrak{S}$ 便分解为所有互不相交的等价类之并集.

注意，这一分解的特点在于：同一子集合中任两元素互相等价，不同子集合中任两元素互不等价. 因此，在每个等价类中任意取出一个元素,则都可以成为这个等价类的代表元素,而且将这些元素(每个集合中取出一个)合并为另一个子集合，则这个子集

合具有性质:

(1) 任两元素互不等价;

(2) 集合 $\mathfrak{S}$ 中任一元素必和这个子集合中某个元素等价.

这个子集合称为**代表元集**.

由上面的讨论可以看出来，这个代表元集的选取，即每个等价类中的代表元素的选取,具有极大的任意性. 事实上,等价类中任一元素都有资格充当代表元素. 所以在具体的集合中，如果定义好等价关系,我们便不满足于随便取出一个元素为代表元素,而要求取出的元素在某种意义下具有最简单的形式. 为了说明这点,下面详细讨论矩阵的情形. 在 § 4.4 中我们引进了

定义 两个 $n\times m$ 矩阵 A 和 B 称为**相抵的**，如果存在一个 n 阶非异方阵 P,及一个 m 阶非异方阵 Q,使得

$$B=PAQ.$$

由定义出发,我们证明了

引理 3 记 M_{nm} 为所有 $n\times m$ 矩阵构成的集合，则 M_{nm} 中矩阵的相抵关系为等价关系.

定理 2 任取 $n\times m$ 矩阵 A. 记 $\operatorname{rank}(A)=r$. 则存在 n 阶非异方阵 P 及 m 阶非异方阵 Q,使得

$$PAQ=\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}=A_r^{(n,m)}$$

由这个定理,所以我们在每个等价类中取如此简单的元素 A_r 为代表元素. 由于在相抵下秩不改变. 所以同一个等价类中任意元素的秩都相同，且不同等价类中的元素的秩决不相同。因此每个等价类由秩完全决定. 所以我们可以记作 $\mathfrak{S}_r$,其中代表元素为 A_r. 其好处在于 $n\times m$ 矩阵 A_r 由秩 r 完全决定,且最简单. 又不难证明

定理 3 $M_{nm}=\bigcup_{r=1}^{\min(n,m)}\mathfrak{S}_r$ 为两两不相交的等价类 $\mathfrak{S}_r$ 之并集.

从这个例子也可以看出，我们在等价类中所寻找的代表元素具有这样的特点：它由一批数量来确定，而且这批数量是同一等价类中每个元素都具有的共性，即这批数量在等价关系下是不变的.所以其中每个量都称为在等价关系下的**不变量**. 不变量的全体称为**全系不变量**. 例如，在相抵关系下，矩阵的秩就是全系不变量.要注意的是：单单不变量并不能区分等价类，即不同的等价类中的元素可以具有相同的不变量，但是不可能具有相同的全系不变量，即全系不变量是能完全区分等价类的一批数量.

对矩阵或方阵集合而言，在其中引进某种等价关系后，具有最简形式的代表元素便称为**标准形**，而矩阵标准形理论的目的就是建立寻找标准形的方法.

1. 方阵的情形

记 $M_n(C)$ 为所有 n 阶复方阵构成的集合. 在其中引进关系 $A\to PAf(P)$，其中函数 $f(P)$ 是将 n 阶非异方阵 P 用转置运算、共轭运算和取逆运算作各种可能的组合. 这样可以导出八种关系：$A\to$

$$PAP,\quad PAP',\quad PA\overline{P},\quad PA\overline{P}'$$

$$PAP^{-1},PA(P^{-1})',\ PA(\overline{P^{-1}}),\ PA(\overline{P^{-1}})'.$$

引理 4 上面八种关系中，四种关系

$$A\to PAP',\ A\to PA\overline{P}',\ A\to PAP^{-1},\ A\to PA\overline{P}^{-1}$$

为等价关系，它们分别称为相合、复相合、相似、复相似. 余下四种关系

$$A\to PAP,\ A\to PA\overline{P},\ A\to PA(P^{-1})',\quad A\to P\overline{A(P^{-1})}'$$

都不是等价关系. 它们都适合反身性和对称性，但不适合传递性.

矩阵论的一个重要部分是研究上面四种等价关系，求出它们的标准形，决定全系不变量.

对 n 阶非异方阵 P 加上限制，我们又可以得到一批新的等价关系. 一般条件为: 设 P 为酉方阵或复正交方阵.

设 $P=U$ 为酉方阵，即 $U\bar{U}'=\bar{U}'U=I^{(n)}$，则等价关系 $A\to UA\bar{U}'$ 和 $A\to UAU^{-1}$ 相同，又 $A\to UA\bar{U}^{-1}$ 和 $A\to UAU'$ 相同. 所以有两种等价关系: $A\to UAU'$, $A\to UAU^{-1}$. 前者称为**酉相合**，后者称为**酉相似**.

设 $P=O$ 为复正交方阵，即 $OO'=O'O=I^{(n)}$. 则等价关系 $A\to OAO'$ 和 $A\to OAO^{-1}$ 相同，又 $A\to OA\bar{O}^{-1}$ 和 $A\to OA\bar{O}'$ 相同，所以有两种等价关系: $A\to OAO^{-1}$，$A\to OA\bar{O}'$. 前者称为**复正交相似**，后者称为**复正交复相合**.

记 $M_n(R)$ 为所有 n 阶实方阵构成的集合. 则只有两种等价关系

$$A\to PAP',\quad A\to PAP^{-1},$$

其中 P 为 n 阶非异实方阵. 前者称为**实相合**，后者称为**实相似**.

由于酉方阵及复正交方阵限制在实的情形，都是实正交方阵. 所以对 n 阶非异实方阵 P 加上条件为实正交方阵，改记为 $P=O$，即有 $OO'=O'O=I^{(n)}$，因此等价关系 $A\to OAO'$ 和 $A\to OAO^{-1}$ 相同. 所以只有一种等价关系: $A\to OAO'$，称为**实正交相似**.

2. 矩阵的情形.

设 $M_{nm}(C)$ 为所有 $n\times m$ 复矩阵构成的集合. 我们只考虑两种等价关系: $A\to PAQ$，$A\to UAV$，其中 P,Q 分别为 n 阶及 m 阶复非异方阵，U, V 分别为 n 阶及 m 阶酉方阵. 前者称为**相抵**，后者称为**酉相抵**.

设 $M_{nm}(R)$ 为所有 $n\times m$ 实矩阵构成的集合. 我们也只考虑两种等价关系: $A\to PAQ$，$A\to O_1AO_2$，其中 P，Q 分别为 n 阶及 m 阶实非异方阵，O_1，O_2 分别为 n 阶及 m 阶实正交方阵. 前者称为**实相抵**，后者称为**实正交相抵**.

上述各种等价关系下的标准形理论，构成本书的主要内容.

作为重要应用，下面利用等价关系引进线性空间的商空间概念. 设 $\mathfrak{L}$ 为 n 维线性空间，$\mathfrak{L}_1$ 为 m 维子空间. 我们引进一种关系，称为"模 $\mathfrak{L}_1$ 相等"如下：任取 $\alpha,\beta\in\mathfrak{L}$，如果 $\alpha-\beta\in\mathfrak{L}_1$，则称 α 和 β 模 $\mathfrak{L}_1$ 相等，记作

$$\alpha\equiv\beta \pmod{\mathfrak{L}_1}.$$

显然，它是等价关系. 事实上，由 $\alpha-\alpha=0\in\mathfrak{L}_1$ 可知 $\alpha\equiv\alpha(\mathrm{mod}\,\mathfrak{L}_1)$，即反身性成立. 再设 $\alpha-\beta\in\mathfrak{L}_1$，即 $\beta-\alpha=-(\alpha-\beta)\in\mathfrak{L}_1$，所以由 $\alpha\equiv\beta(\mathrm{mod}\,\mathfrak{L}_1)$ 可知 $\beta\equiv\alpha(\mathrm{mod}\,\mathfrak{L}_1)$，即对称性成立. 最后，设 $\alpha-\beta\in\mathfrak{L}_1$，$\beta-\gamma\in\mathfrak{L}_1$，则 $\alpha-\gamma=(\alpha-\beta)+(\beta-\gamma)\in\mathfrak{L}_1$，所以由 $\alpha\equiv\beta(\mathrm{mod}\,\mathfrak{L}_1)$，$\beta\equiv\gamma(\mathrm{mod}\,\mathfrak{L}_1)$ 有 $\alpha\equiv\gamma(\mathrm{mod}\,\mathfrak{L}_1)$，即传递性成立，至此证明了断言.

在 $\mathfrak{L}$ 中按模 $\mathfrak{L}_1$ 相等引进等价关系后，$\mathfrak{L}$ 可分为互不相交的等价类之并集，等价类全体构成的集合记作 $\mathfrak{L}/\mathfrak{L}_1$. 在 $\mathfrak{L}/\mathfrak{L}_1$ 中引进加法及纯量积如下：在等价类中任取一个代表元素 α，我们记 $\tilde{\alpha}$ 为此等价类. 任取 $\alpha,\beta\in\mathfrak{L}$，$a$ 为数，定义

$$\tilde{\alpha}+\tilde{\beta}=\widetilde{(\alpha+\beta)},\quad a(\tilde{\alpha})=\widetilde{(a\alpha)}.$$

这样就定义了加法和纯量积. 但是定义依赖于代表元素之选取，所以为了证明上述定义只是等价类之间的运算，需要证明它们实际上与代表元素的选取无关. 事实上，任取 $\xi\in\tilde{\alpha}$，$\eta\in\tilde{\beta}$，由定义可知 $\xi-\alpha,\eta-\beta\in\mathfrak{L}_1$. 即存在 $\sigma,\tau\in\mathfrak{L}_1$ 使得

$$\xi=\alpha+\sigma,\quad \eta=\beta+\tau.$$

于是 $\xi+\eta=(\alpha+\beta)+(\sigma+\tau)$. 由 $\sigma+\tau\in\mathfrak{L}_1$，这证明了 $\xi+\eta\in\widetilde{(\alpha+\beta)}$，因此 $\tilde{\xi}=\tilde{\alpha}$，$\tilde{\eta}=\tilde{\beta}$，$\widetilde{(\xi+\eta)}=\widetilde{(\alpha+\beta)}$. 这证明了 $\tilde{\xi}+\tilde{\eta}=\tilde{\alpha}+\tilde{\beta}=\widetilde{(\alpha+\beta)}=\widetilde{(\xi+\eta)}$. 于是加法运算与代表元选取无关. 再任取 $\xi\in\tilde{\alpha}$，a 为数. 则由 $\xi-\alpha\in\mathfrak{L}_1$，所以存在 $\sigma\in\mathfrak{L}_1$，且 $\xi=$

$\alpha+\sigma$. 于是 $a\xi=a\alpha+a\sigma$. 但是 $a\sigma\in\mathfrak{L}_1$, 所以 $a\xi\in(\widetilde{a\alpha})$. 这证明了 $a\tilde{\xi}=a\tilde{\alpha}=(\widetilde{a\alpha})=(\widetilde{a\xi})$. 即纯量运算也与代表元素选取无关.

由线性空间的定义及直接验证可知集合 $\mathfrak{L}/\mathfrak{L}_1$ 在上述加法及纯量积下构成线性空间,称为 $\mathfrak{L}$ 模 $\mathfrak{L}_1$ 之**商空间**.

为方便起见,我们用符号 $\alpha+\mathfrak{L}_1$ 表示以 α 为代表元素的等价类. 实际上,它有很直观的含义. 因为每个和元素 α 等价之元素必形如 $\alpha+\sigma$,其中 $\sigma\in\mathfrak{L}_1$,反之,任取 $\sigma\in\mathfrak{L}_1$,则 α 和 $\alpha+\sigma$ 必等价. 用新的符号,则加法及纯量积可定义为

$$(\alpha+\mathfrak{L}_1)+(\beta+\mathfrak{L}_1)=(\alpha+\beta)+\mathfrak{L}_1,\quad \forall\alpha,\beta\in\mathfrak{L};$$

$$a(\alpha+\mathfrak{L}_1)=a\alpha+\mathfrak{L}_1\qquad \forall\alpha\in\mathfrak{L},a\text{ 为数}.$$

很自然地,我们可以建立线性空间 $\mathfrak{L}$ 到商空间 $\mathfrak{L}/\mathfrak{L}_1$ 上的映射 $\pi:\alpha\to\alpha+\mathfrak{L}_1$. 显然它的核 $\pi^{-1}(0)=\mathfrak{L}_1$. 上面加法及纯量积定义告诉我们有

$$\pi(\alpha)+\pi(\beta)=\pi(\alpha+\beta),\quad a\pi(\alpha)=\pi(a\alpha).$$

这证明了 π 为 $\mathfrak{L}$ 到 $\mathfrak{L}/\mathfrak{L}_1$ 上的线性映射,称为**自然同态**. 由定理 7.1.4,有维数公式

$$\begin{aligned}\dim\mathfrak{L}_1&=\dim\pi(\mathfrak{L}_1)+\dim\pi^{-1}(0)\\&=\dim\mathfrak{L}/\mathfrak{L}_1+\dim\mathfrak{L}_1.\end{aligned}$$

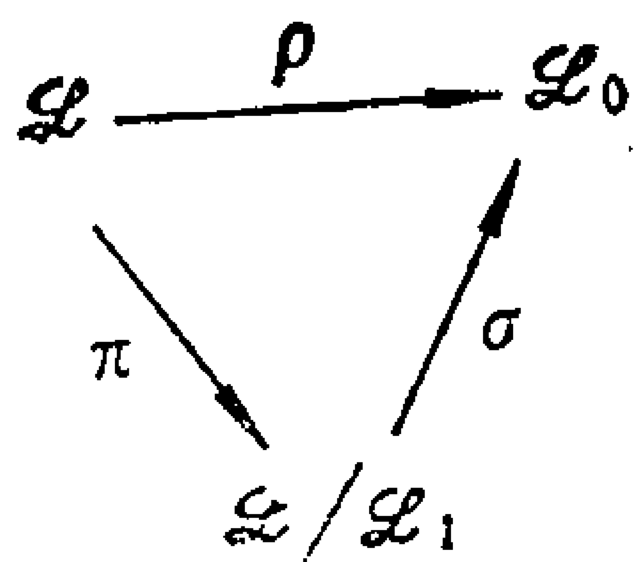
即有

$$\dim\mathfrak{L}/\mathfrak{L}_1=\dim\mathfrak{L}-\dim\mathfrak{L}_1.$$

最后,我们给出**同态基本定理**: 设 $\rho:\mathfrak{L}\to\mathfrak{L}_0$ 为线性空间 $\mathfrak{L}$ 到线性空间 $\mathfrak{L}_0$ 上的线性映射. 同态核 $\rho^{-1}(0)=\mathfrak{L}_1$ 为 $\mathfrak{L}$ 的子空间. 记 $\pi:\mathfrak{L}\to\mathfrak{L}/\mathfrak{L}_1$ 为自然同态,则存在线性空间 $\mathfrak{L}/\mathfrak{L}_1$ 到线性空间 $\mathfrak{L}_0$ 上的同构映射 σ,使得

$$\rho=\sigma\circ\pi,$$

即下图



可交换．证明如下：任取 $\alpha\in\mathfrak{L}$，则 $\pi(\alpha)=\alpha+\mathfrak{L}_1$．建立 $\mathfrak{L}/\mathfrak{L}_1$ 到 $\mathfrak{L}_0$ 上之对应 σ：

$$\alpha+\mathfrak{L}_0\to\rho(\alpha),\quad \forall\alpha\in\mathfrak{L}$$

我们来证明 σ 为 $\mathfrak{L}/\mathfrak{L}_1$ 到 $\mathfrak{L}_0$ 上的线性同构，且有 $\rho=\sigma\circ\pi$．事实上，任取 $\alpha\in\mathfrak{L}$，则

$$(\sigma\circ\pi)(\alpha)=\sigma(\pi(\alpha))=\sigma(\alpha+\mathfrak{L}_0)=\rho(\alpha).$$

这证明了 $\sigma\circ\pi=\rho$．再证 σ 一一．今若 $\alpha+\mathfrak{L}_1=\beta+\mathfrak{L}_1$，于是 $\alpha-\beta\in\mathfrak{L}_1$，所以由 $\rho(\mathfrak{L}_1)=0$ 有 $\rho(\alpha-\beta)=0$，即 $\rho(\alpha)=\rho(\beta)$．这证明了 σ 为单值的．反之，任取 α，$\beta\in\mathfrak{L}$，且 $\rho(\alpha)=\rho(\rho)$，所以 $\rho(\alpha-\beta)=0$，即 $\alpha-\beta\in\mathfrak{L}_1$，这证明了 $\alpha+\mathfrak{L}_1=\beta+\mathfrak{L}_1$，即 σ 为一一映射．最后证 σ 为线性同构．事实上，任取 $\alpha,\beta\in\mathfrak{L}$，$a,b$ 为数，则

$$\begin{aligned}&\sigma(a(\alpha+\mathfrak{L}_1)+b(\beta+\mathfrak{L}_1))=\sigma((a\alpha+b\beta)+\mathfrak{L}_1)\\ &=\rho(a\alpha+b\beta)=a\rho(\alpha)+b\rho(\beta)\\ &=a\sigma(\alpha+\mathfrak{L}_1)+b\sigma(\beta+\mathfrak{L}_1).\end{aligned}$$

这证明了 σ 为 $\mathfrak{L}/\mathfrak{L}_1$ 到 $\mathfrak{L}_0$ 上的线性同构．至此证明了断言．

习题

1．下列关系是否为等价关系：

(1) 方阵的交换关系；

(2) 给定区间 $[a,b]$ 上一连续函数 $f(x)$．区间 $[a,b]$ 中的实数 x_1 和 x_2 称为有关系的，如果 $f(x_1)=f(x_2)$．

2．设集合 $\mathfrak{C}$ 中有一种关系"$\to$"，它具有对称性和传递性，试证明下列推理的不正确性：

"今由对称性，从 $a\to b$ 可知 $b\to a$，于是有 $a\to b,b\to a$．再由传递性可知 $a\to a$．所以反身性成立"．

3．为什么说等价关系是数的相等概念的推广．

习题答案与提示

习题 1.1

1. $-5x^7+13x^6+27x^5-130x^4+75x^3+266x^2-440x+197$.

2. $f(x)=x^4+x^3+x-1$.

3. $u(x)=x^{-m}[1-(1-x^m)^n]$, $v(x)=(1-x)^{-n}(1-x^m)^n$.

7. 将 f_1 与 f_2 作辗转相除，将 (f_1,f_2) 与 f_3 再作辗转相除. 将辗转相除的每一步都对行列式的列做相应的变化. 最后得到的行列式取作

$$\begin{vmatrix} a(f_1,f_2,f_3) & 0 & 0 \\ 0 & a^{-1} & 0 \\ 0 & 0 & 1 \end{vmatrix}=(f_1,f_2,f_3).$$

再将整个过程倒过来，便证明了存在性. 另法. 存在多项式 u_1, u_2, v, v_3, 使得 $\begin{vmatrix} f_1 & f_2 \\ -u_2 & u_1 \end{vmatrix}=(f_1,f_2)$, $v\begin{vmatrix} f_1 & f_2 \\ -u_2 & u_1 \end{vmatrix}+v_3f_3=(f_1,f_2,f_3)$. 将所求行列式按第 3 列作 Laplace 展开，再与上式比较.

习题 1.2

1. 先判定 f 和 g 的次数. $f(x)=\pm\left(x^2+\frac{1}{2}x+1\right)$, $g(x)=\frac{\sqrt{5}}{2}x$ 或 $g(x)=-\frac{\sqrt{5}}{2}x$.

2. $m=2^{k+1}-1$.

3. 双方都乘 $1-x^2$，判断何时 $(1+x,1-x^n)=1$. 答案为 n 是奇数.

4. 记 $b=D(x)$. 对 $f(x)=x^n$ 证明结论，再对一般多项式证明.

5. 在 u_1,u_2 中至少有一个是一次多项式时，无妨设 u_1 是. 令 $y=u_1(x)$，故无妨设 $u_1=x$. 再证 $(u_3-u_1)\mid n_x$, $(u_3-u_2)\mid u_1$. 分 $u_2=0$ 及 $u_2\neq 0$ 两种情形讨论，以便证明 $u_2(0)=u_3(0)=0$.

6. 由 $f'(x)=\sum\limits_{i=1}^{n}\frac{f(x)}{x-\alpha_i}$，可证 $g_k(x)=x^{k+1}f'(x)-\sum\limits_{i=1}^{n}\frac{x^{k+1}-\alpha_i^{k+1}}{x-\alpha_i}f(x)$.

再将 $f(x)=x^n-\sigma_1x^{n-1}+\cdots+(-1)^n\sigma_n$ 代入，比较两边 x^n 项之系数，便得 Newton 公式.

7. 考虑多项式的标准分解式. 用到 $\max(a,\min(b,c))=\min(\max(a,b),\max(a,c))$ 对实数 a,b,c 都成立.

习题 1.3

1. (ii) 用反证法证明 x^6-x^2+1 模 2 不可约. (iii) 用反证法. 设能分解为 $f_1(x)f_2(x)$，再证 $f_1(x)=-f_2(x)$.

2. 乘以 $x^5\mp1$.

3. 用 Eisenstein 判别法.

习题 1.4

1. 考虑 $(p(x),f(x))$.

3. 记 $G(x)=f_1(x)f_2(x)-f_2(x_0)f_1(x_0)$. 用归纳法证 $G^{(t)}(x_0)=0,t=0,1,\cdots,k,G^{(k+1)}(x_0)\neq0$.

4. 记 $f(x)=(x+1)^n-x^n-1$. 化简 (f,f') 为 $(x^{n-1}-1,(x+1)^{n-1}-1)$. 答案是 $n\neq6m+1,m=1,2,\cdots$.

6. 注意到 $f(1)=f'(1)=\cdots=f^{(k)}(1)=0$，并利用归纳法.

7 考察 $\dfrac{f}{(f,f')}$ 的次数.

8. 记 $F(x)=f(x)+f'(x)+\cdots+f^{(n)}(x)$. 去证 $F(x)$ 的任一极小值非负，且 $F(\pm\infty)\geqslant0$.

9. 利用 5 次单位根.

11. $2x^2-2x+4$.

12. 设 $f(x)=a_0\prod\limits_{j=1}^{s}(x-\alpha_j)^{e_j}$，其中 $\alpha_1,\cdots,\alpha_s$ 为不同实数. 又设 $f^2=g^2+h^2$，其中 g,h 是次数不等的实多项式，由 $g(\alpha_j)^2+h(\alpha_j)^2=0$ 可导出矛盾. 反之，若 f 有非实复根 α，则 $f=f_1\cdot(x-\alpha)(x-\bar{\alpha})$，其中 f_1 为实多项式. 取 $g=f_1\cdot\mathrm{Re}((x-\alpha)^2),h=f_1\cdot\mathrm{I_m}((x-\alpha)^2)$，也可导出矛盾.

13. 用反证法. 考虑 $g(x)=x^nf\left(\dfrac{1}{x}\right)$. 用到根与系数的关系.

14. 设 $f(x)=a_0(x-x_1)^{r_1}\cdots(x-x_k)^{r_k}(x-x_{k+1})^{r_{k+1}}\cdots(x-x_l)^{r_l}(x-\bar{x}_{k+1})^{r_{k+1}}\cdots(x-\bar{x}_l)^{r_l}$，其中 $x_1<x_2<\cdots<x_k$ 是实数，$x_{k+1},\cdots,x_l$ 是 $l-k$ 个不同非实复数. 证明 $a_0\geqslant0,r_1,\cdots,r_k$ 都是偶数. 故可记 $f(x)=\prod\limits_{j=1}^{m_1}(x-x_j)^2\cdot$

$\prod_{i=1}^{m_2}(x^2+p_ix+q_i)$，其中 $p_i^2-4q_i<0$．再用归纳法证明 $\prod_{i=1}^{m_2}(x^2+p_ix+q^i)=\Phi_1^2+\Phi_2^2$，其中 Φ_1,Φ_2 是实多项式．

15. 用反证法．求 f',f''，证明 $f(x)$至少有三个根不同，再分情形导出矛盾．

16. 令 $x=2\cos\theta$，去求出 $p_n(x)-x$ 的 2^n 个不同的实根，并证明这就是它的全部复根．

17. 记题中多项式为 $f(x)$．利用 $f(x)$ 系数的对称性，可设 $f(x)=(x^2+cx+1)(x^2+dx+1)$，其中 c,d 是实数，$c^2\geqslant4$．推出 $a^2+b^2=(c+d)^2+(cd+2)^2$．求得 $\min(a^2+b^2)=\frac{4}{5}$．

18. 化为考虑 $F(x)=1-\frac{p(x)}{a_0x^n}$的正根，利用导数与函数增减性的关系．

19. 用反证法．

20. 记两多项式的另一个根分别是 β,δ．可证 $\beta-\delta=r-p$．故只须证 $r=p$．若 $r\neq p$，可得 $\alpha=\frac{q-s}{r-p}$及 $r-p|1$，与 α 不是整数矛盾．

21. 首先证明，若题中多项式可约，必可表为一个整多项式的平方．答案是$(a,b,c)\neq\pm(3,2,1)$．

22. 设 $f(x)=f_1(x)f_2(x)$．先证明 $f_1(a_j)$与$f_2(a_j)$ 同为 1 或者同为 -1 不妨设同为 1，再表出 $f_1(x)-1$ 与 $f_2(x)-1$，推出矛盾．

23. 先设 $p(k)=1,p(l)=-1$，去证明$|k-l|\leqslant2$．再进一步具体讨论满足上述的 k 与 l 的个数．

24. 将第一个方程乘以 a^2，第二个方程乘以 a，由$\sum_{k=1}^{5}[a^2kx_k-2ak^3x_k+k^5x_k]=0$ 推得 $a=0,1,4,9,16,25$．

25. 设 α,β 是 $f(x)$的任两不同根．若 $m=\alpha+\beta$ 是有理数，可证 $g(x)=f(m-x)$ 也是有理系数不可约多项式．再证明 $g(x)=-f(x)$，推出$\frac{m}{2}$是 $f(x)$的根，与 $f(x)$不可约矛盾．

26. $f_0(x)=a_0$．用归纳法证明 $f_n(x)=\sum_{j=0}^{n}a_j\binom{n}{j}x^j$，其中用到

$\binom{n}{i}\binom{i}{j}=\binom{n}{j}\binom{n-j}{i-j}$ 及 $\sum_{k=0}^{l}(-1)^k\binom{l}{k}2^k=(-1)^l$. 最后直接计算验证题中等式.

27. $-x^2+1$.

28. $f(x)=a\prod_{k=0}^{25}(x-k)$, a 是常数.

29. 当 $\deg(f)\leqslant 0$ 时 $f(x)=0$ 或 $f(x)=e$, 其中 $e^{n-1}=1$. 当 $\deg(f)>0$ 时, 对任意复数 a, 去证 $f(a)=a^n$, 所以 $f(x)=x^n$.

30. $f(x)=x$.

习题 2.2

3. (i) 0. (ii) $4\sin^4\theta$.

5. 有 $(n^2)!$ 种不同的方阵, 再利用行列式的性质, 将其中行列式的值相等的看作一个.

习题 2.3

1. $a_1a_2a_3a_4\sum_{i=1}^{4}\frac{1}{a_i}$.

3. 在 $\sum_{j=1}^{n}\det A_j$ 中把 $\det A_j$ 按第 j 行展开, 再重新组合.

4. 证明所有的 A_{ij} 都等于 A_{11}.

5. 把等式左端各列都加到第 1 列上, 再提出 $(n-a)$.

习题 2.4

1. 0.

2. 从最后一列起, 把每列乘以 x 加到它的前一列上. 答案是

$$\sum_{j=1}^{n}a_jx^{n-j}.$$

3. 把第 i 行提出 $\sin\theta_i$, $i=1,\cdots,n$. 再把 $\frac{\sin mx}{\sin x}$ 表为 $\cos x$ 的多项式. 还用到 Vandermonde 行列式. 答案是 $2^{\frac{n(n-1)}{2}}\prod_{i=1}^{n}\sin\theta_i\prod_{1\leqslant s<t\leqslant n}(\cos\theta_t-\cos\theta_s)$.

4. $a^{n-2}(a^2-1)$.

5. 对 n 用归纳法，将右边按第 1 行展开.

6. 0.

7. 添一行一列，第 1 行第 1 列元素为 1，第 1 行其余元素为 -1，第 1 列其余元素为 0. 答案是 $\prod_{1\le s<i\le n}(x_i-x_s)\prod_{k=0}^{n}y_k\sum_{j=0}^{n}\frac{(-1)^j}{y_j}\sum_{1\le i_1<\cdots<i_{n-j}\le n}x_{i_1}\cdots x_{i_{n-j}}$.

8. 按第 1 行展开求递推公式. 答案是 $\prod_{i=1}^{n}(a_ia_{2n+1-i}-b_ib_{2n+1-i})$.

9. 把 a_n 写成 $(a_n-b)+b$，按最后一列拆成两个行列式，可得一递推公式. 再类似地按最后一行拆开，可得另一递推公式. 两递推公式联立求解. 答案是 $\frac{1}{b-c}\left[b\prod_{i=1}^{n}(a_i-c)-c\prod_{i=1}^{n}(a_i-b)\right]$.

10. 将 0 写成 $1-1$，拆成两个行列式之和，再利用上题结果. 答案是

$$\frac{1}{b-c}\left[\prod_{i=1}^{n-1}(a_i-b)-\prod_{i=1}^{n-1}(a_i-c)\right].$$

11. 将 b_n 写成 $(b_n-a_n)+a_n$ 去求递推公式. 答案是

$$\prod_{i=1}^{n}(b_i-a_i)\left(1+\sum_{j=1}^{n}\frac{a_j}{b_j-a_j}\right).$$

12. 将最后一行乘以 (-1) 加到其余各行上去，提出因子后再将最后一列乘以 (-1) 加到其余各列上去，再提出因子，求得递推公式. 答案是

$$\prod_{1\le k<i\le n}(a_i-a_k)(b_i-b_k)\prod_{j,l=1}^{n}\frac{1}{a_j+b_l}.$$

13. 从最后一行起每行减去它的上一行. 用到公式 $\binom{m}{l}+\binom{m}{l+1}=\binom{m+1}{l+1}$. 答案是 1.

14. 用 § 2.4 例 1 的方法. 答案是 $\frac{1}{1!2!\cdots(n-1)!}\prod_{1\le j<i\le n}(m_i-m_j)$.

15. 0.

16. 化为下三角形. 利用组合公式推出主对角线上元素的规律.

17. 由最后一行起将每行加到前一行上，再由最后一列起每列减去它的前一列. 求出递推公式. 答案是$(x-n)^{n+1}$.

18. $2^{n-1}\prod_{i=1}^{n} a_i$.

19. 添加一行一列使成 Vandermonde 行列式，添加的一行是 $1 x_1 \cdots x_1^n$. 考虑新行列式的结果中 x^{k+1} 的系数. 答案是

$$\sum_{1\leq j_1<\cdots<j_{n-(k+1)}\leq n} a_{j_1}\cdots a_{j_{n-(k+1)}} \prod_{1\leq t<s\leq n}(a_s-a_t).$$

20. 对 n 用归纳法

21. 用归纳法.

23. 当 n 为偶数时，把行列式的第 1 行乘以$\frac{a_{2j}}{a_{12}}$加到第 j 行上去，把第 2 行乘以$\left(-\frac{a_{1j}}{a_{12}}\right)$加到第 j 行上去，$j=3,\cdots,n$. 再对阶数用归纳法.

习题 2.5

1. 令 $x_1=x^2, x_2=x$，将 y 看作常数，化为线性方程组. 答案是 $(-2,0)$; $(-3,1)$; $(1,3)$; $(0,4)$.

2. (i) $x_j=\frac{1}{(a_j-a_1)\cdots(a_j-a_{j-1})(a_j-a_{j+1})\cdots(a_j-a_n)}\cdot\sum_{k=1}^{n}(-1)^{n+k}b_k$ $\varphi_{k,j}$，其中 $\varphi_{k,j}=\sum_{\substack{1\leq i_1<\cdots<i_{n-k}\leq n\\ i_1,\cdots,i_{n-k}\neq j}} a_{i_1}a_{i_2}\cdots a_{i_{n-k}}$, $j=1,2,\cdots,n$. (ii) $x_k=(-1)^{n+k}$ $\sum_{i=1}^{n}\frac{b_i\varphi_{k,i}}{(a_i-a_1)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_n)}$，其中 $\varphi_{k,i}=\sum_{\substack{1\leq t_1<\cdots<t_{n-k}\leq n\\ t_1,\cdots,t_{n-k}\neq i}}$ $a_{t_1}a_{t_2}\cdots a_{t_{n-k}}$, $k=1,2,\cdots,n$.

习题 3.1

1. 先证明$\frac{\partial\sigma_k}{\partial x_j}=\sigma_{k-1}-x_j\sigma_{k-2}+x_j^2\sigma_{k-3}-\cdots+(-1)^{k-2}x_j^{k-2}\sigma_1+(-1)^{k-1}x_j^{k-1}$. 答案是：若 $f(x_1,\cdots,x_n)=\varphi(\sigma_1,\cdots,\sigma_n)$，则$\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}=\sum_{k=1}^{n}$

$(n-k+1)\sigma_{k-1}\frac{\partial\varphi}{\partial\sigma_k}$.

2. 先证明辅助命题:"若 $f(x_1,\cdots,x_n)$ 在未知数的偶排列下不变,而在奇排列下反号,则 $f(x_1,\cdots,x_n)=h(x_1,\cdots,x_n)V(x_1,\cdots,x_n)$ 其中 h 是对称多项式,$V(x_1,\cdots,x_n)$ 是 Vandermonde 行列式."再把 $f(x)$ 拆成两个多项式的和,使前一个是对称多项式,后一个符合辅助命题的条件.

3. 先证明任意多项式 $f(x_1,\cdots,x_n)$ 都可表为 $g_0,g_1,\cdots,g_{n-1}$ 的多项式,再证明 $f(x_2,x_3,\cdots,x_n,x_1)=\varepsilon^{-(i_1+2i_2+\cdots+(n-1)i_{n-1})}f(x_1,\cdots,x_n)$. 用到本原单位根的性质.

4. 用反证法. 用到行列式的定义.

5. 可记 n 次齐次多项式 $f_n(x)=x^n+a_1x^{n-1}y+\cdots+a_ny^n$,记 $g(x)=x^n+a_1x^{n-1}+\cdots+a_n$,则 $f_n(x,y)=y^ng\left(\frac{x}{y}\right)$. 当 $n>1$ 时,取 $x=y=-\frac{1}{2}$,$z=1$,由条件可证 $g(-1)=0$,即 $(x+1)|g(x)$. 由此推出 $f_n(x,y)=(x+y)f_{n-1}(x,y)$,其中 $f_{n-1}(x,y)$ 是 $n-1$ 次齐次多项式,并可验它仍然适合条件,所以可对 n 用归纳法. 注意到 $f_1(x,y)=x-2y$,故 $f_n(x,y)=(x+y)^{n-1}(x-2y)$.

习题 3.2

1. (i)用定理 3.2.1. (ii)设 $f(x)=a_0(x-x_1)\cdots(x-x_n)$,再用判别式的定义. (iii)记 $\deg(g)=m$,可设 $g(x)-x_j$ 的 m 个根为 $y_1^{(j)},\cdots,y_m^{(j)}$,$j=1,\cdots,n$. 用到定理 3.2.3,定理 3.2.1 及 1(i)的结果.

2. 用 1(iii)的结果. 记 $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n$,答案是

$$(-1)^{\frac{mn(m-1)}{2}}m^{mn}a_n^{m-1}(\Delta(f))^m.$$

习题 4.1

3. (ii) $\begin{pmatrix}\cos n\theta & \sin n\theta\\ -\sin n\theta & \cos n\theta\end{pmatrix}$. (iii)把矩阵化为上题的形式. (iv)用归纳法. 答案是

$$\begin{pmatrix}\lambda^p & C_p^1\lambda^{p-1} & C_p^2\lambda^{p-2} & \cdots & C_p^{n-1}\lambda^{p-n+1}\\ & \lambda^p & C_p^1\lambda^{p-1} & \cdots & C_p^{n-2}\lambda^{p-n+2}\\ & \ddots & \ddots & \ddots & \vdots\\ & & & & C_p^1\lambda^{p-1}\\ & & & \ddots & \lambda^p\end{pmatrix}.$$

4. 用定理 2.3.1 的推论 2.

7. 用归纳法

8. 利用 $AE_{ij}=E_{ij}A, i, j=1,\cdots,n$.

9. 将与 $A=\mathrm{diag}(\lambda_1 I_1,\cdots,\lambda_s I_s)$ 可交换的方阵 B 按与 A 相同的方式分块. 比较 $AB=BA$ 的两边.

习题 4.2

1. 用行列式性质可证 $\det\begin{pmatrix} A & 0 \\ -I & B \end{pmatrix}=\det\begin{pmatrix} A & AB \\ -I & 0 \end{pmatrix}$.

2. 用定理 4.2.1 及定理 2.3.1 的推论 2.

3. (i)分别左乘、右乘一个适当的对角矩阵. 用到定理 4.2.1. 答案为 $\prod_{j=1}^{n} f(\varepsilon^j)$,其中 $f(x)=a_0+\mu^{\frac{1}{n}}a_1x+\cdots+\mu^{\frac{n-1}{n}}a_{n-1}x^{n-1}$, $\varepsilon=\cos\frac{2\pi}{n}+\sqrt{-1}\sin\frac{2\pi}{n}$. (ii)用 Vandermonde 行列式及定理 4.2.1. 答案为 $\prod_{1\le s<t\le n}(x_t-x_s)^2$.

4. (i)分 $\det A$ 是否为 0 讨论. (ii) 用定理 4.2.1 及定理 2.3.1 的推论 2.

7. 用定理 4.2.2.

8. 乘以转置矩阵的行列式.

9. 用定理 4.2.2, 注意到双重和号可交换.

10. 取 $P=\begin{pmatrix} \lambda I^{(n)} & -A \\ -B & I^{(m)} \end{pmatrix}$.

习题 4.3

2. 用 $\mathrm{tr}AB=\mathrm{tr}BA$.

3. 用习题 4.2 第 10 题的结果. 当 $1-\lambda B'A\neq 0$ 时 $I^{(n)}-\lambda AB'$ 有逆, 逆为 $I^{(n)}+\frac{\lambda}{1-\lambda B'A}AB'$.

4. 对 m 用归纳法.

5. 用逆方阵的定义.

6. 前者的逆方阵为 $\frac{1}{n}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega^{-1} & \cdots & \omega^{-(n-1)} \\ \vdots & \vdots & & \vdots \\ 1 & \omega^{-(n-1)} & \cdots & \omega^{-(n-1)^2} \end{pmatrix}$,后者的逆方阵为

$$\begin{pmatrix} (I-PQ)^{-1} & -P(I-QP)^{-1} \\ -Q(I-PQ)^{-1} & (I-QP)^{-1} \end{pmatrix}.$$

7. 用习题 4.2 第 10 题的结果. $(I-BA)^{-1}=I+B(I-AB)^{-1}A$.

8. 用"摄动法".

9. 右乘 $\begin{pmatrix} 0 & I^{(m)} \\ I^{(n)} & -B \end{pmatrix}$.

习题 4.4

3. 设 $A=P\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q$, 将 $QB'P$ 适当分块.

4. 设 $A=P\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q$, 推出 $B=(P')^{-1}\begin{pmatrix} B_1^{(r)} & 0 \\ B_3 & B_4 \end{pmatrix}Q$, 其中 $B_1=B_1'$.

5. 用到定理 4.4.5 和定理 4.2.2.

7. (i) 凑出 $\begin{pmatrix} I-A & -A \\ I & I \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & I-A \end{pmatrix}\begin{pmatrix} I-A & I \\ -A & I \end{pmatrix}=\begin{pmatrix} A-A^2 & 0 \\ 0 & I \end{pmatrix}$, 并证明 $\begin{pmatrix} I-A & -A \\ I & I \end{pmatrix}$ 及 $\begin{pmatrix} I-A & I \\ -A & I \end{pmatrix}$ 均非异. (ii) 类似(i).

8. 利用矩阵相抵下的标准形.

9. 先证"矩阵删去一行,秩最多减少 1".

10. 用打洞技巧.

11. 最大秩是 $n-1$.

12. (i) 用 $AA^*=\det A\cdot I$. (ii) 用到习题 4.2 第 2 题. (iii) 用秩的定义. (iv) 分 A 是否非异讨论.

13. 利用 $\operatorname{rank}(A+B)\leqslant\operatorname{rank}(A)+\operatorname{rank}(B)$.

14. 用矩阵相抵下的标准形.

15. 设 $A=P\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}Q, B=M\begin{pmatrix} I^{(n-r)} & 0 \\ 0 & 0 \end{pmatrix}N$, 则 $X=Q^{-1}\begin{pmatrix} 0^{(r,n-r)} & X_2 \\ X_3 & X_4 \end{pmatrix}M^{-1}$.

16. 对 p 用归纳法.

17. 利用分块矩阵及相抵下的标准形.

18. 用到定理 4.4.5.

19. 先证明"若 $\operatorname{rank}(B)=\operatorname{rank}(AB)$, 则 $\operatorname{rank}(BC)=\operatorname{rank}(ABC)$".

习题 5.1

1. 若 $I^{(n)}-A$ 奇异, 则 $(I-A)x=0$ 必有非零解, 从而导出矛盾. 后一

论断的证明用到定理 5.1.2.

2. 设 $a_i\neq 0$，则通解为 $x_{t_j}=\frac{a_i a_j}{a_i^2}x_{ti}, i,j=1,2,\cdots,n$.

3. 用"矩阵的秩不小于它的子矩阵的秩". 不是必要条件.

4. 用矩阵相抵下的标准形. 当 $r(A)=m$ 时解唯一.

5. 若不然，取 $Ax=0$ 的非零解，设其分量中绝对值最大者为第 i_0 个分量，由第 i_0 个方程推出矛盾.

6. 充分性用到"同解的线性方程组系数矩阵同秩".

习题 5.2

1. 将$\det\begin{pmatrix} a_{j1} & a_{j2} & \cdots & a_{jn+1} \\ a_{11} & a_{12} & \cdots & a_{1n+1} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn+1} \end{pmatrix}$按第 1 行展开，$j=1,2,\cdots,n$.

习题 5.3

2. 用特征根及特征向量的定义. $\det f(A)=\prod_{j=1}^{n}f(\lambda_j)$.

3. 取 A 的属于 λ_0 的特征向量；利用 Cauchy 不等式.

4. $\alpha\bar{\alpha}'$ 的特征根是 0 和 $\bar{\alpha}'\alpha$.

5. 不对，因为没有区分数零和零矩阵.

6. 用圆盘定理.

7. 用定理 5.3.2. 最后一命题还用到 Newton 公式.

8. 用定理 5.3.2.

9. 用定理 5.3.1.

10. 记 $f(t)=\det(aI+tk), t\in[0,1]$，用到连续函数的性质.

11. 条件即$A'\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$.

习题 6.1

2. 在棋盘上取定坐标系后利用第 1 题的结果.

3. 用到定理 5.2.1.

6. 对方阵的不同特征根的个数用归纳法.

7. 先设 $\beta=\sum_{i=1}^{n}a\alpha_i$，再证明所有的系数都是0.

8. 在后一问中不妨设 $c_1c_2\cdots c_m\neq 0$，可推出 $\beta_m=-\frac{1}{c_m}\sum_{i=1}^{m-1}c_i\beta_i$ 及 $\beta_m=-\frac{1}{b_m}\sum_{i=1}^{m-1}b_i\beta_i$.

习题 6.2

1. 用后一条件证明该向量组线性无关.

2. 写出 $\{\alpha_1,\cdots,\alpha_n\}$ 到 $\{\beta_1,\cdots,\beta_n\}$ 的基变换的矩阵 (a_{ij})，这个矩阵是非异的，从而可推出必有 $12\cdots n$ 的排列 $i_1i_2\cdots i_n$，使得 $a_{1i_1}a_{2i_2}\cdots a_{ni_n}\neq 0$.

习题 6.4

1. (i) 是；$n-1$ 维. (ii) 不是；线性生成的子空间为 n 维. (iii) 同(ii). (iV)不是；线性生成的子空间为 n 维，一组基为 $(2,1,\cdots,1,-1,\cdots,-1)',\cdots,(1,\cdots,1,2,-1,\cdots,-1)',(1,\cdots,1,-2,-1,\cdots,-1)',\cdots,(1,\cdots,1,-1,\cdots,-1,-2)'$.

3. 考虑极大线性无关部分组.

4. (V) 三个子空间 $\mathfrak{L}_1,\mathfrak{L}_2,\mathfrak{L}_3$ 中有两个之间有包含关系时，就相等. (vi)同(v).

5. 不唯一.

6. (1)⇒(2)：先证明 $\dim\sum_{j=1}^{t}\mathfrak{L}_j\leqslant\sum_{j=1}^{t}\dim\mathfrak{L}_j$，再用定理 6.4.4. (2)⇒(3)：注意到 $\sum_{s=1}^{j-1}\mathfrak{L}_s\subset(\mathfrak{L}_1+\cdots+\mathfrak{L}_{j-1}+\mathfrak{L}_{j+1}+\cdots+\mathfrak{L}_t)$. (3)⇒(1)：用归纳法.

7. (i) 对 t 用归纳法. 设 $\alpha\notin\mathfrak{L}_j$，$1\leqslant j\leqslant t-1$，$\alpha\in\mathfrak{L}_t$，$\beta\notin\mathfrak{L}_t$，考察 $k_j\alpha+\beta,1\leqslant j\leqslant t$，其中 k_j 两两不同. 可证 $k_j\alpha+\beta\notin\mathfrak{L}_t$，且它们中任意两个都不属于同一个 $\mathfrak{L}_j,1\leqslant j\leqslant t-1$. (ii)用(i).

8. 利用第 7 题.

9. 考虑矩阵行向量组的极大线性无关部分组.

习题 6.5

1. 记 $\mathfrak{L}_2=\{x\in V_m|Bx=0\}$，可得 $\mathfrak{L}_2\subset\mathscr{L}_0$. 取 $\mathfrak{L}_2$的一组基 $\alpha_1,\cdots,\alpha_{m-r(B)}$，扩充成 $\mathfrak{L}_0$ 的一组基 $\alpha_1,\cdots,\alpha_{m-r(B)},\alpha_{m-r(B)+1},\cdots,\alpha_{m-r(AB)}$. 再证明 $B\alpha_{m-r(B)+1},\cdots,B\alpha_{m-r(AB)}$线性无关，从而得到 $\dim\mathfrak{L}_1=\mathrm{rank}(B)$

$-\text{rank}(AB)$. 然后记 $W_0=\{x\in V_n \mid ABx=0\}$，$W_1=\{y\in V_n \mid y=Bx, x\in W_0\}$，$H_0=\{x\in V_n \mid ABCx=0\}$，$H_1=\{y\in V_n \mid y=BCx, x\in H_0\}$，通过证明$H_1\subset W_1$来证明最后那个不等式.

2. 考虑基础解系.

3. 考虑 $ABCx=0$ 与 $BCx=0$ 的基础解系.

4. 前一问的答案为$\begin{pmatrix}A\\B\end{pmatrix}x=0$. 后一问的答案为 $Cx=0$，其中 C 的行向量组是A的行向量的生成空间与B的行向量组的生成空间的交空间的一组基.

5. 注意到B的列向量组的生成空间就是 $Ax=0$ 的解空间，其中 $x\in V_n$.

习题 7.1

2. 对 $\mathfrak{L}$ 中的任意向量 α、β，构造适当的线性变换 $\mathscr{A}_1$，$\mathscr{A}_2$，以便从 $\mathscr{A}(\mathscr{A}_1+\mathscr{A}_2)=\mathscr{A}\mathscr{A}_1+\mathscr{A}\mathscr{A}_2$ 推出 $\mathscr{A}(\alpha+\beta)=\mathscr{A}(\alpha)+\mathscr{A}(\beta)$. 再由此推出对任意整数 p，有 $\mathscr{A}(p\alpha)=p\mathscr{A}(\alpha)$，及对任意分数$\frac{p}{q}$，有 $\mathscr{A}\left(\frac{p}{q}\alpha\right)=\frac{p}{q}\mathscr{A}(\alpha)$. 最后取极限，证明对任意实数 c，有 $\mathscr{A}(c\alpha)=c\mathscr{A}(\alpha)$.

3. 当 $A=0, D=-C$时，$\mathscr{A}$ 在基$\{E_{ij}, i,j=1,\cdots,n\}$下的方阵表示为

$$\begin{pmatrix} a_{11}I-C' & a_{12}I & \cdots & a_{1n}I \\ a_{21}I & a_{22}I-C' & \cdots & a_{2n}I \\ \vdots & \vdots & & \vdots \\ a_{n1}I & a_{n2}I & \cdots & a_{nn}I-C' \end{pmatrix}.$$

4. 是线性变换.

5. $\mathfrak{L}$ 是$(n+1)^2$维的. 用定义可证 τ 为 $\mathfrak{L}$ 上线性变换. τ 在基$\left\{\frac{1}{j!k!}x^jy^k, 0\leqslant j, k\leqslant n\right\}$下的方阵表示为 $4\begin{pmatrix} D & & & \\ & I+D & & \\ & & \ddots & \\ & & & nI+D \end{pmatrix}$，其中

$$D=\begin{pmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & n+1 \end{pmatrix}.$$

习题 7.2

1. 前一命题的必要性，可设 α，$\mathscr{A}(\alpha),\cdots,\mathscr{A}^{n-1}(\alpha)$ 为 $\mathfrak{L}$ 的一组基，写出 $\mathscr{A}$ 在这组基下的方阵表示，然后求特征向量. 后一命题的必要性可采用反证法. 两者的充分性可以类似地证明如下. 先设 $\mathscr{A}$ 的特征多项式为 $f(\lambda)=(\lambda-\lambda_1)^{e_1}\cdots(\lambda-\lambda_s)^{e_s}$，其中 $\lambda_1,\cdots,\lambda_s$ 两两不同，去证明 $\mathscr{A}$ 的 Jordan 标准形必形为 $J=\mathrm{diag}(\lambda_1I_1+N_1,\cdots,\lambda_sI_s+N_s)$，其中 I_i 是 e_i 阶单位方阵. 然后记 $f(\lambda)=\lambda^n-a_1\lambda^{n-1}-\cdots-a_n$，构造方阵 $B=\begin{pmatrix} a_1 & 1 & & \\ a_2 & 0 & \ddots & \\ \vdots & \vdots & \ddots & 1 \\ a_n & 0 & \cdots & 0 \end{pmatrix}$，去证明 B 与 J 相似，因而 B 是 $\mathscr{A}$ 在某组基下的方阵表示.

3. (i) 对空间的维数用归纳法. 先证明"任一线性变换的特征子空间，是与之可交换的线性变换的不变子空间." (ii) 对方阵的阶数用归纳法. (iii) 对空间的维数用归纳法. 把空间写成其中一个线性变换的特征子空间的直接和，再把其余线性变换限制在这些特征子空间上，写出 Jordan 标准形，利用 Jordan 标准形的唯一性去证明.

4. 先证明 A、B 可同时相似于上三角形.

5. 满足 $(\lambda_jI-A)x=0$ 的所有 $x\in V_n$ 是属于特征根 λ_j 的全体特征向量，它们都在特征根 λ_j 的根子空间中.

6. 当 $\dim(\mathscr{A}(\mathfrak{L}_1\cap\mathfrak{L}_2))=\dim(\mathscr{A}(\mathfrak{L}_1)\cap\mathscr{A}(\mathfrak{L}_2))$ 时.

7. 由 $u(\lambda)g(\lambda)+v(\lambda)h(\lambda)=1$ 得 $u(\mathscr{A})g(\mathscr{A})+v(\mathscr{A})h(\mathscr{A})=id$. 并用 Hamilton-Cayley 定理. 再设 $f(\lambda)=(\lambda-\lambda_1)^{e_1}\cdots(\lambda-\lambda_s)^{e_s}$，其中 $\lambda_1,\cdots,\lambda_s$ 两两不同，则 $(\lambda-\lambda_1)^{e_1},\cdots,(\lambda-\lambda_s)^{e_s}$ 两两互素.

8. $\sum_{i=1}^{n}\alpha_i$ 是循环向量.

9. 用 Jordan 标准形.

10. 转换成线性变换的命题去做. 先证明"任一线性变换的根子空间，是与之可交换的线性变换的不变子空间". 再证明可有一个线性变换存在低维的奇数维实根子空间，然后把其余的线性变换限制在这个根子空间上. 用到 3(i) 题的结论.

11. 利用 $u(x)f(x)+v(x)g(x)=1$.

12. 证明 $\mathscr{A},\mathscr{B}$ 的 Jordan 标准形都是对角形，再用 3(iii) 题的结论.

13. 用定理 7.1.4.

14. 设$\sigma\neq0$,去证σ为$\mathfrak{L}_1$到$\mathfrak{L}_2$上的同构. 用$\sigma(\mathfrak{L}_1)$是任一$\mathfrak{L}$的不变子空间证明$\sigma(\mathscr{L}_1)=\mathfrak{L}_2$; 用$\ker(\sigma)$是任一$\mathscr{A}$的不变子空间证明$\ker(\sigma)=0$. 最后一问可设$\alpha\neq0,\sigma(\alpha)=a\alpha$,去证$(\sigma-a(id))(\mathfrak{L}_1)$是任一$\mathscr{A}$的不变子空间.

15. 先证明与A对应的线性变换$\mathscr{A}$的象空间与核构成空间直接和.再分别取基,拼成整个空间的基.

16. (1)⇒(2): $\mathrm{tr}A_i=\mathrm{rank}(A_i),i=1,\cdots,m$. (2)⇒(3): 在$A_1,\cdots,A_m$中任取$s$个加起来记为$A$, 余下的加起来记为$B$.可征$\mathrm{rank}(A)+\mathrm{rank})B=n$. 再无妨设$A$为Jordan标准形去证明$\begin{pmatrix}I&0\\0&0\end{pmatrix}$, $B=\begin{pmatrix}0&0\\0&I\end{pmatrix}$, 从而$AB=0$.最后取$A=A_i;A=A_i+A_j$,可证$A_i=\begin{pmatrix}I&0\\0&0\end{pmatrix}$;$A_j=\begin{pmatrix}0&0\\0&D_j\end{pmatrix}$.

另法: 构造分块方阵$\mathrm{diag}(A_1,\cdots,A_m)$.

习题 8.1

1. n元二次齐次函数是二次型.

2. 用到齐次线性方程组理论.

3. (i)$\forall\alpha,\beta\in\mathfrak{M}$,考虑$f(\alpha+\beta,\alpha+\beta)=0$. (ii)先证明$\mathfrak{M}_1\cap\mathfrak{M}_2^{\perp}=\mathfrak{M}_1\cap\mathfrak{P}_2^{\perp}$,再证明$\mathfrak{M}_1\cap\mathfrak{M}_2\subset\mathfrak{P}_2^{\perp}$, 然后用互相包含的方法证明$\mathfrak{M}_1\cap\mathfrak{M}_2{}_2^{\perp}=(\mathfrak{M}_1\cap\mathfrak{M}_2)+(\mathfrak{P}_1\cap\mathfrak{P}_2^{\perp})$,最后证明是直接和.

习题 8.2

1. 根据外积的定义及排列的性质.

3. 前一问用多重线性函数的定义和s次齐次多项式的定义. 后一问的存在性只须对任一s次n元单项式$F(x)=x_1^{i_1}\cdots x_n^{i_n}$证明.记$\xi_j=(x_{1j}\cdots x_{nj})'$, $j=1,\cdots,s$, $f_0(\xi_1,\cdots,\xi_s)=x_{11}\cdots x_{1i_1}x_{2i_1+1}\cdots x_{2i_1+i_2}\cdots x_{ni_1+\cdots+i_{(n-1)}+1}\cdots x_{ns}$.
再令$f(\xi_1,\cdots,\xi_s)=\dfrac{1}{s!}\sum\limits_{\binom{1\cdots s}{j_1\cdots j_s}}f_0(\xi_{j_1}\cdots\xi_{j_s})$,则它是$\mathfrak{L}$上$s$重对称线性函数,且$f(x,\cdots,x)=F(x)$. 唯一性先化为证明"当$f(\xi_1,\cdots,\xi_s)$为$s$重对称线性函数,$f(x,\cdots,x)=0$时, 必有$f(\xi_1,\cdots,\xi_s)=0$". 再引入参数$t$,由$f(x+t\xi_1,\cdots,x+t\xi_1)=0$证明$f(\xi_1,x,\cdots,x)=0$. 用此证明$f(\xi_1,\xi_2,x,\cdots,x)=0$. 继续做下去,最后证明$f(\xi_1,\cdots,\xi_s)=0$.

习题 9.1

1. $\forall\alpha,\beta\in\mathfrak{L}$, 数$a$, 证明$(\mathscr{A}(a\alpha+\beta)-a\mathscr{A}(\alpha)-\mathscr{A}(\beta)$, $\mathscr{A}(a\alpha+\beta)-$

$a\mathscr{A}(\alpha)-\mathscr{A}(\beta))=0$，从而证明 $\mathscr{A}$ 是线性变换.

2. $\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$.

3. 利用实正交方阵的定义，对方阵的阶数用归纳法.

4. $\det(A+B)=\det A\det(B^{-1}+A^{-1})\det B=-\det(A+B)$.

5. 设 $OA\alpha=\lambda\alpha$，$\alpha\neq0$，将此式两边取共轭转置后再与原式相乘，可得 $\bar{\alpha}'A^2\alpha=|\lambda|^2\bar{\alpha}'\alpha$.

6. 不妨设所考虑的子方阵位于左上角，取 $A=\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$，用上题的结论.

7. 设 n 阶实正交方阵 $O=(a_{ij})$，则命题等价于"可用一些题中形状的矩阵左乘或右乘 O 得到单位矩阵". 不妨设 $a_{11}\neq0$，用"乘以题中形状矩阵"的办法把 $a_{21},\cdots,a_{n1}$ 都变成 0. 由于乘得的仍是实正交矩阵，故第 1 行及第 1 列中除左上角的元素为 ±1 外其余全是 0，右下角是 $n-1$ 阶实正交矩阵. 再对阶数用归纳法.

9. 用线性方程组理论. 当 $G(\alpha_1,\cdots,\alpha_n)=G(\beta_1,\cdots,\beta_n)$ 时，构造线性变换 $\mathscr{A}$，使 $\mathscr{A}(\alpha_i)=\beta_i$，$i=1,\cdots,n$，去证 $\mathscr{A}$ 保持向量的长度不变.

10. 先证明存在置换方阵 P，使 $PA=\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$，其中 A_1 的 r 个行向量线性无关. 再用定理 9.1.7.

11. 设 $O\alpha=\lambda_0\alpha$，$\alpha\neq0$. 两边取共轭转置，再相乘. 当 λ_0 为实数时，两边取转置，再相乘.

12. O 必以 1 为一个特征根. 设另两个特征根为 x_1,x_2，考察特征多项式 $f(\lambda)=(\lambda-1)(\lambda-x_2)(\lambda-x_3)$.

13. 设 $K\alpha=\lambda_0\alpha$，$\alpha\neq0$. 两边取共轭转置，再右乘 α. 设 $K(\beta+\sqrt{-1}\,\gamma)=\sqrt{-1}\,b(\beta+\sqrt{-1}\gamma)$，$b\neq0$. 得 $K\beta=-b\gamma$，$K\gamma=b\beta$. 两式分别左乘 β'，再利用"一行一列的矩阵等于自身的转置".

14. 用归纳法分别证 $r\geqslant n+1$ 和 $r\leqslant n+1$，从而得 $r=n+1$.

15. (ii) 对 s 用归纳法，设 $(\tilde{\beta}_0,\beta_j)=0$，$1\leqslant j\leqslant s-1$. 若 $(\tilde{\beta}_0,\beta_s)=0$，则考虑 $\beta_0=\tilde{\beta}_0+\varepsilon\beta_s$，其中 $\varepsilon>0$. 利用连续性，取适当的 ε. (iii)类似(ii).

16. 注意到 $\mathfrak{L}_1^{\perp},\cdots,\mathfrak{L}_r^{\perp}$ 也都是真子空间，再利用习题 6.4 第 7(i)题结论，对维数用归纳法.

17. $(A+I)^{-1}+(A'+I)^{-1}=I$ 等价于 $A'+I+A+I=(A+I)(A'+I)$.

习题 9.2

1. 对方阵的阶数用归纳法. 据习题 7.2 第 3 题知，这些实规范方阵一定有公共的复特征向量 α. 再分 α 是实向量和非实向量两种情形讨论. 仿照定理 9.2.1 的证明.

2. 对 $\sum_{i=1}^{n} s_j^2 = 0$ 两边取迹.

4. s_1 是唯一的. 参见定理 10.2.4 的证明.

5. 利用定理 9.2.5(9.2.6)，并且注意到一个矩阵左右分别乘同一个置换方阵时，主子式仍然变为主子式.

6. 用定理 9.2.2. 并用到 AB 与 BA 有相同的特征根.

7. K 为实斜对称方阵等价于 $(I-O')(I+O)=(I+O')(O-I)$.

8. 有实正交方阵 O，使 $OAO^{-1}=\begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$，其中 A_{11} 是对角元素都为 1 的下三角方阵，A_{22} 是对角元素都为 0 的下三角方阵. 再证明 $A_{11}=I$，$A_{22}=0$. 记 $A_1=A_{21}$，取 $P=\begin{pmatrix} I^{(r)} & 0 \\ -A_1 & I \end{pmatrix}O$.

9. 实对称方阵 s 幂等时，$a_{jj}=\sum_{k=1}^{n} a_{jk}{}^2$.

10. 在不变子空间 $\mathfrak{L}_1$ 中取一组标准正交基 $\alpha_1, \cdots, \alpha_r$，拼成整个空间的一组标准正交基. 实规范变换在这组基下的方阵表示为 $A=\begin{pmatrix} A_1 & A_2 \\ 0 & A_4 \end{pmatrix}$. 利用迹的性质去证 $A_2=0$.

11. $n-\text{rank}(A+B)=n-\text{rank}(I+A^{-1}B)$. 考虑 $A^{-1}B$ 的标准形中一的个数.

12. 参见第 8 题.

习题 9.3

1. (i)先证明 e 是 s 的属于 λ_n 的特征向量，再将 $\frac{1}{\sqrt{n}}e$ 扩充成空间的一组标准正交基，并分别考察 s 及 $I-\frac{1}{n}ee'$ 对这组基的作用，以证明 $s-\lambda_{n-1}\left(I-\frac{1}{n}ee'\right)$ 的特征根均不小于 0. (ii) 用到“半定正实对称方阵的对角元均不小于 0”.

习题 10.1

1. 对 s 的阶数用归纳法. 利用矩阵的打洞技巧.

2. 必要且充分条件是：秩为 1，或秩为 2 而符号差为 0. 按两个实线性函数是否成比例分情况讨论.

3. 设 $s=\begin{pmatrix}0^{(k)} & B\\ B' & D^{(n-k)}\end{pmatrix}$. 用实相合关系把 s 化为

$$\begin{pmatrix}0^{(k)} & \frac{1}{2}I^{(k)} & 0\\ \frac{1}{2}I^{(k)} & 0 & 0\\ 0 & 0 & D^{(n-2k)}\end{pmatrix}.$$ 其中用到 B 在相抵下的标准形.

习题 10.2

2. 记 $\beta=(b_1,\cdots,b_n)'$, $x=(x_1,\ \cdots,\ x_n)'$, 则 $f(x_1,\ \cdots,\ x_n)=(x'\ \ 1)\begin{pmatrix}A & \beta\\ \beta' & a\end{pmatrix}\begin{pmatrix}x\\ 1\end{pmatrix}$. 对$\begin{pmatrix}A & \beta\\ \beta' & a\end{pmatrix}$用打洞技巧. 答案是：$f$ 有极小值 $a-\beta'A^{-1}\beta$.

4. 实对称方阵半定正的必要且充分条件是所有特征根都非负.

5. 由 $s_1-aI\geqslant 0, s_2-cI\geqslant 0$ 推出 $(s_1+s_2)-(a+c)I\geqslant 0$, 再推出 s_1+s_2 的特征根不小于 $a+c$.

6. s_1 与 s_2 可同时实正交相似于对角形.

7. 用定理 10.2.1 及定理 9.1.7. 后一问化为证明：Q_1Q^{-1} 既是实正交方阵又是上三角方阵.

8. 用 $AA'=A'A$ 及定理 10.2.4 中的唯一性.

9. 记 $S=\begin{pmatrix}S_1 & \alpha\\ \alpha' & a\end{pmatrix}$, 用打洞技巧去证明 S 实相合于$\begin{pmatrix}S_1 & 0\\ 0 & 0\end{pmatrix}$.

10. 对 n 用归纳法. 参见第 9 题.

11. 用第 10 题结论. 注意到 $|\det A|=(\det(A'A))^{\frac{1}{2}}$ 而 $A'A$ 定正.

12. 不妨设 S 定正，则存在非异实方阵 P，使 $P'SP=I$, $P'TP=\mathrm{diag}(a_1,\cdots,a_n)$.

13. 不妨设 S 定正. 记 $S=\begin{pmatrix}S_1^{(r)} & B\\ B' & S_2\end{pmatrix}$, 用打洞技巧，再利用第12 题结论. 最后一问可对方阵的阶数用归纳法.

14. 首先，有不全为零的 $n\times 1$ 实矩阵 β_1, β_2, 使 $(S_1+\sqrt{-1}S_2)(\beta_1+$

$\sqrt{-1}\beta_2)=0$. 其次，将实部与虚部分开，再利用 $S_1\geqslant 0$的条件，证明$\beta_i' S_1\beta_i=0$，从而证明 $S_1\beta_i=0$ 及 $S_2\beta_i=0, i=1,2$.

16. 参见引理 9.1.1 的证明.

17. 将函数变形为 $f(x_1,\cdots, x_n)=\frac{1}{2}\left(x_1^2+\sum_{i=1}^{n-1}(x_i-x_{i+1})^2+x_n^2\right)$，极小值为 0. 在 $x_n=1$ 的限制下极小值为$\frac{n+1}{2n}$.

18. 有实正交方阵 O，使 $O'SO=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$，$\lambda_i>0, i=1,\cdots,n$. 记 $Y_k=O'X_kO$，去证 Y_k 都是对角矩阵. 记 $Y_k=\mathrm{diag}(a_1^{(k)},\cdots,a_n^{(k)})$，便转化为证明序列 $a_j^{(k+1)}=\frac{1}{2}\left(a_j^{(k)}+\frac{\lambda_j}{a_j^{(k)}}\right)$，$(j=1,\cdots,n)$有极限，且极限大于 0. 再对 $X_{k+1}=\frac{1}{2}(X_k+SX_k^{-1})$ 取极限，便得 $B^2=S$. 最后一问用归纳法去证 $X_kP=PX_k$.

19. 用打洞技巧并证明 $B'A^{-1}B$ 可逆.

20. 注意 $\mathrm{rank}(A'A)=\mathrm{rank}(A)$.

21. 用实正交相似下的标准形化为证明 $\lambda_i+\frac{1}{\lambda_i}-2\geqslant 0$.

22. 用第 21 题结论.

23. 不妨设 S 为实相合下的标准形 I. 再将$\frac{1}{2}(A-A')$ 化为实正交相似下的标准形.

24. 先构造$\begin{pmatrix} I & Y \\ X' & I\end{pmatrix}\begin{pmatrix} I & -Y \\ -X' & I\end{pmatrix}$以证明辅助命题："若 X, Y 是 n 阶实矩阵，$I-X'X\geqslant 0$，$I-Y'Y\geqslant 0$，则有$(\det(I-X'Y))^2\geqslant\det(I-X'X)\det(I-Y'Y)$". 再设 $A_1\geqslant 0$，$A_1^2=A$，记 $B=A_1O$，利用辅助命题去证 $|\det(I-AO)|\geqslant\det(I-A)$. 最后证明$I-AO$ 的实特征根都不小于 0，从而 $\det(I-AO)\geqslant 0$.

25. 不妨设 $k=1$. 这是 $n-1$ 个变元的二次型. 用到习题 2.4 第 5 题的结论.

26. 当 $AB=0$ 时，可证 A，B 可同时实正交相似于对角形. 反之，当 $\lambda^n\det(\lambda I-A-B)=\det(\lambda I-A)\det(\lambda I-B)$时，先把问题化为只须证明

A和B的非零特征根个数之和等于n的情形. 再不妨设$A=\begin{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_r \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix}$,

且有实正交方阵O,使$OBO'=\begin{pmatrix} 0 & 0 \\ 0 & \begin{pmatrix} \lambda_{r+1} & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \end{pmatrix}$, 其中 $\lambda_1\cdots\lambda_n\neq 0$. 然后

将O分块为$\begin{pmatrix} X^{(r)} & Y \\ U & V \end{pmatrix}$,去证明$U=0$,从而推出$Y=0$. 最后验证$AB=0$.

习题 11.1

2. 用定义. 注意到$\overline{\omega}^k=\omega^{n-k}$.

3. 记$U=\begin{pmatrix} |a|e^{\sqrt{-1}\varphi_1} & |b|e^{\sqrt{-1}\varphi_2} \\ |c|e^{\sqrt{-1}\varphi_3} & |d|e^{\sqrt{-1}\varphi_4} \end{pmatrix}$, 设法找到合适的$\theta_1,\theta_2,\theta_3,\theta_4$,使

$\begin{pmatrix} e^{-\sqrt{-1}\theta_1} & 0 \\ 0 & e^{-\sqrt{-1}\theta_2} \end{pmatrix}U\begin{pmatrix} e^{-\sqrt{-1}\theta_3} & 0 \\ 0 & e^{\sqrt{-1}\theta_4} \end{pmatrix}$为实方阵,从而就是实正交方阵.

4. $\{E_{ij}, i,j=1,\cdots,n\}$是该内积下的一组标准正交基.

习题 11.2

2. 先证明A,B,C可同时酉相似于对角形. 再证明由于特征根两两不同,它们的特征子空间全是一维的. 最后证明A,B,C的分别属于λ_j,a_j,b_j的特征子空间是同一子空间.

3. 参见习题9.2第7题.

4. 记$\alpha=\beta+\sqrt{-1}\gamma$,其中$\beta,\gamma\in R^n$. 先证明存在$n$阶实正交方阵$O_1$,使$(O_1\beta)'=(a\ 0\ \cdots\ 0)$. 再记$(O_1\gamma)'=(b\ \ \gamma_1')$,去证明存在$n$阶实正交方阵$O_2$,使$(O_2\alpha)=(a\ 0\ \cdots\ 0)'+\sqrt{-1}(b\ c\ 0\ \cdots\ 0)'$. 然后对$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$考虑实正交相抵下的标准形,并注意到二阶实正交方阵总能写为$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$与$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$中的一种.

5. λ_0^{-1}是$U^{-1}A$的特征根. 再利用习题9.1第5题在复数范围内的推广.

6. 设$Ax=\lambda x$, $x=(x_1,\cdots,x_n)'\neq 0$. 记$x_{j_0}$为$x_j$中模最大者,记$A=$

(a_{ij})，将 $\lambda x_{j_0}=\sum\limits_{j=1}^{n}a_{j_0j}x_j$两边取模，去证$|\lambda|\leqslant na$. 再推出 $b\bar{x}'x=\bar{x}'\frac{A+\bar{A}'}{2}x$，将右边写成坐标形式，适当放大，去证$|b|\leqslant nh$. 类似可证$|c|\leqslant nk$.

7. 用到定理 11.2.1. 还用到"相似矩阵有相同的迹和相同的特征根".

8. (i) 记 $S=A+\sqrt{-1}B$，其中 A,B 为实方阵. 证明A，B实对称，且可交换，从而可同时实正交相似于对角形. (ii) 注意 $\mathrm{diag}(e^{\sqrt{-1}\theta_1},\cdots,e^{\sqrt{-1}\theta_n})=\mathrm{diag}(e^{\frac{\sqrt{-1}}{2}\theta_1},\cdots,e^{\frac{\sqrt{-1}}{2}\theta_n})\cdot\mathrm{diag}(e^{\frac{\sqrt{-1}}{2}\theta_1},\cdots,e^{\frac{\sqrt{-1}}{2}\theta_n})$. (iii) 参见第十三章. 取 S_1 实正交相似于 $\mathrm{diag}(\theta_1,\cdots,\theta_n)$. (iV) 对任一酉方阵$U$，证明$U'U=O\mathrm{diag}(e^{\sqrt{-1}\theta_1},\cdots,e^{\sqrt{-1}\theta_n})O'$，其中$O$是实正交方阵. 记 $S=O\mathrm{diag}(e^{\frac{\sqrt{-1}}{2}\theta_1},\cdots,e^{\frac{\sqrt{-1}}{2}\theta_n})O'$，可推出$U=\bar{U}S\cdot S$，再证明$\bar{U}S$是实方阵. 分解不唯一. (V) 利用(iV)和(i).

9. (i) 类似第8(i)题. (ii) 由(i)有实正交方阵Z，使 $O=Z\mathrm{diag}\left(\begin{pmatrix}a_1 & \sqrt{-1}b_1\\ -\sqrt{-1}b_1 & a_1\end{pmatrix},\cdots,\begin{pmatrix}a_s & \sqrt{-1}b_s\\ -\sqrt{-1}b_s & a_s\end{pmatrix},1,\cdots,1,-1,\cdots,-1\right)Z'$. 再取 $\varepsilon_j=\mathrm{sgn}a_j$，则有实数 x_j，使 $\varepsilon_ja_j=\mathrm{ch}x_j$，$\varepsilon_jb_j=\mathrm{sh}x_j$. 然后记 $K_j=\begin{pmatrix}0 & x_j\\ -x_j & 0\end{pmatrix}$，$T_0=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -\sqrt{-1}\\ -\sqrt{-1} & 1\end{pmatrix}$，则$T_0$是酉方阵，且 $T_0\begin{pmatrix}x_j & 0\\ 0 & -x_j\end{pmatrix}T_0'=\sqrt{-1}K_j$. 再据第十三章有

$$e^{\sqrt{-1}k_j}=\varepsilon_j\begin{pmatrix}a_j & \sqrt{-1}b_j\\ -\sqrt{-1}b_j & a_j\end{pmatrix}.$$

最后令 $K=Z\mathrm{diag}(K_1,\cdots,K_s,0,\cdots,0)Z'$，$A=Z\mathrm{diag}(\varepsilon_1,\varepsilon_1,\cdots,\varepsilon_s,\varepsilon_s,1,\cdots,1,-1,\cdots,-1,)Z'$，验证 $O=Ae^{\sqrt{-1}k}$. 分解不唯一. (iii) 先证明$\bar{O}'O$是定正的正交 Hermite 方阵，再利用(ii). 分解不唯一.

10. (ii) 取 α_1，使 $H\alpha_1=\lambda_1\alpha_1$，$|\alpha_1|=1$；取$\alpha_2=-J\bar{\alpha}_1$. 再在$[\alpha_1,\alpha_2]^\perp$中取 α_3，使 $H\alpha_3=\lambda_2\alpha_3$，$|\alpha_3|=1$；取 $\alpha_4=-J\bar{\alpha}_3$. 类似地取下去，证明 $\alpha_1,\cdots,\alpha_{2n}$ 为标准正交基. 然后令 $U=(\alpha_1\cdots\alpha_{2n})$，去验证结论. (iii) 取 α_1，使 $A\alpha_1=\lambda_1\alpha_1$，$|\alpha_1|=1$；取 $\alpha_2=-J\bar{\alpha}_1$. 用类似(ii)的方法得到标准正交基 $\alpha_1,\cdots,\alpha_{2n}$. 令 $U=(\alpha_1,\cdots,\alpha_n)$. 对$n$用归纳法.

习题 11.3

1. (i) $\operatorname{diag}(1,\cdots,1,-1)$. (ii) $\operatorname{diag}(1,\cdots,1)$.

2. (i) 当 n 为奇数时，为 $\sum_{j=1}^{\frac{n-1}{2}} y_j\bar{y}_j - \sum_{j=\frac{n+1}{2}}^{n-1} y_j\bar{y}_j$；当 n 为偶数时，为 $\sum_{j=1}^{\frac{n}{2}} y_j\bar{y}_j - \sum_{j=\frac{n}{2}+1}^{n} y_j\bar{y}_j$. (ii) $y_1\bar{y}_1 - \sum_{j=2}^{n} y_j\bar{y}_j$.

3. (i) I. (ii) I.

习题 11.4

2. 先证 $A\geqslant 0$，从而有非异实方阵 P，使 $P'AP=\begin{pmatrix} I^{(r)} & 0 \\ 0 & 0 \end{pmatrix}$，其中 $r=\operatorname{rank}(A)$. 再证 $P'BP=\begin{pmatrix} B_1^{(r)} & 0 \\ 0 & 0 \end{pmatrix}$，用到“半定正 Hermite 方阵的二阶主子式均非负”，

3. 第一问用打洞技巧和归纳法. 后两问先证 $\operatorname{Re}(H)$ 为定正实对称方阵，从而有 $P'\operatorname{Re}(H)P=I$，其中 P 实非异. 再考虑 $P'I_m(H)P$ 在实正交相似下的标准形.

4. 先证存在置换方阵 P，使 $PJP=\operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\cdots,\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)$. 再参见习题 **11.2** 第 11 题. 最后一问用 Schmidt 正交化.

5. 利用第 3 题结论.

6. 存在非异方阵 Q，使 $\bar{Q}'H_1Q=I$，再将 $\bar{Q}'H_2Q$ 在酉相似下化为标准形.

7. (i) 用第 6 题结论. (ii) 用到不等式 $1+\prod_{j=1}^{n} a_j^{\frac{1}{n}} \leqslant \prod_{j=1}^{n}(1+a_j)^{\frac{1}{n}}$，$(a_j\geqslant 0)$. 这一不等式可以用对称多项式的知识给出证明，

8. 先证明 $\bar{A}'(A\bar{A}')^{-1}A$ 是幂等的，所以是半定正 Hermite 方阵. 再利用习题 10.2 第 12 题的结论在复数范围内的推广.

9. 不妨设 $I-A\bar{A}'$ 及 $I-B\bar{B}'$ 均定正，构造 $2n$ 阶方阵 $\begin{pmatrix} I & \bar{B}' \\ A & I \end{pmatrix}\begin{pmatrix} I & -\bar{B}' \\ -A & I \end{pmatrix}$，用做矩阵乘法和不做矩阵乘法这两种方法计算它的行列式,

比较其结果.

10. 利用第 6 题结论.

11. 先证明存在非异复方阵 P, Q, 使 $P'BP=I, P'AP=Q\mathrm{diag}(\lambda_1,\cdots,\lambda_n)Q^{-1}$, 其中 $\lambda_i\neq\lambda_j, i\neq j$. 由此得 $Q'P'APQ=Q'Q\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$, 再证明 $Q'Q$ 必为对角矩阵.

12. 充分性可直接验证. 必要性的证明先将 A 写成酉相抵下的标准形 $A=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V$, 其中 $\Lambda=\mathrm{diag}(\rho_1I_1,\cdots,\rho_sI_s)$, $\rho_1>\cdots>\rho_s>0$, 再令 $B=\bar{V}'\begin{pmatrix}B_1 & B_2\\ B_3 & B_4\end{pmatrix}\bar{U}'$. 然后设法把问题化为由 "$\Lambda B_1\bar{B}_1'\Lambda=\bar{B}_1'\Lambda^2B_1$, $\Lambda\bar{B}_1'B_1\Lambda=B_1\Lambda^2\bar{B}_1'$, $B_2=0, B_3=0$" 证明 "$\Lambda^2B_1=B_1\Lambda^2, \bar{B}_1'B_1\Lambda=\Lambda B_1\bar{B}_1'$, $B_2=0, B_3=0$". 再记 $B_1=\begin{pmatrix}B_{11} & \cdots & B_{1s}\\ \vdots & & \vdots\\ B_{s1} & \cdots & B_{ss}\end{pmatrix}$, 去证明 $B_{ij}\neq 0(i\neq j)$.

习题 12.1

2. 用强广义逆矩阵的定义.

3. 设法在 $A(\bar{B}')^+$ 中凑出 AB 的因子来. 后一命题可转化为前一命题.

4. 设 $A=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$, $B=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$, 则 $A^+=\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}$, $B^+=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$, 可验证 $(AB)^+\neq B^+A^+$.

5. 设 $A=(a_1\cdots a_n)'$, 若 $a_1=\cdots=a_n=0$, 则 $A^+=(\overbrace{0\cdots 0}^{n\text{个}})$; 若 $a_1,\cdots,a_n$ 不全为 0, 则 $A^+=\left(\sum\limits_{i=1}^{n} a_i\bar{a}_i\right)^{-1}(\bar{a}_1\cdots\bar{a}_n)=(\bar{A}'A)^{-1}\bar{A}'$. 后者可由前者取转置得到.

6. 利用矩阵相抵下的标准形, 有 $A=P\begin{pmatrix}I^{(r)} & 0\\ 0 & 0\end{pmatrix}Q$, 其中 P, Q 分别是 n, m 阶非异方阵. 记 $P=(BB_1)$, $Q=\begin{pmatrix}C\\ C_1\end{pmatrix}$, 其中 B 有 r 列, C 有 r 行. 则 B, C 满足要求. 再用强广义逆矩阵的等价定义, 可验证 $A^+=\bar{C}'(C\bar{C}')^{-1}\cdot(\bar{B}'B)^{-1}\bar{B}'$. 然后利用 C 行满秩, B 列满秩, 去证最后一个等号.

习题 13.1

1. 利用 λ 矩阵在相抵下的标准形.

2. 参见习题1.1第7题"提示"中方法 I.

习题 13.2

1. 参见习题7.2第3(iii)题.

2. 先证明"初等置换方阵的乘积的逆方阵，就相当于把原来的乘积反序".所以，若 P 是初等置换方阵的乘积，则 PAP^{-1} 就相当于对 A 作 P 所表示的一系列的行的互换，再对其结果作一系列相应的列的互换.

4. 证明 A 的极小多项式无重根.

5. 先证明 N^p 的初等因式都形为 λ 的方幂. 再求出不同方幂的初等因式的个数. 用到 N^p 与 J^l 有相同的核，其中 J 是 N^p 的 Jordan 标准形. 答案是 $\mathrm{diag}(J_1,\cdots,J_{p-h},J_{p-h+1},\cdots,J_p)$，其中 $h=n-\left[\frac{n}{p}\right]p$，$J_1=\cdots=J_{p-h}=$

$$\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}^{([\frac{n}{p}])}, J_{p-h+1}=\cdots=J_p=\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}^{([\frac{n}{p}]+1)},$$

这里方括号表示"整数部分".

6. 前一问先证明可在相似下把 A 的左上角元素化为0，并保持条件不变，然后对方阵的阶数用归纳法. 后一问先证明不妨假设 A 的对角元素都是0，再取 n 阶方阵 $X=\begin{pmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & n \end{pmatrix}$ 及 $Y=\left(\frac{a_{ij}}{i-j}\right)$（其中 $i=j$ 时元素为0），这里 $A=(a_{ij})$.

7. 化为考虑 A 的 Jordan 标准形.

8. 用行列式因式.

10. 先化为只须证明 A 是 Jordan 块的情形. 再证明 A 的各级行列式因式或者是1，或者以1为根. 从而证明 A 的 $n-1$ 级行列式因式是1，这里 n 是 A 的阶数.

11. 先证明 A_{ii} 可同时相似于对角形，再证明可不妨假设 $A_{ii}=E_{ii}$，$i=1,\cdots,n$. 然后证 $A_{ij}=a_{ij}E_{ij}$，可证 $a_{ij}a_{jq}=a_{iq}$，$i,j,q=1,\cdots,n$. 最后把问题化为只须证明存在 n 阶非异方阵 $P=(p_{rs})$，使 $p_{rs}=0(r\neq s)$，$p_{ii}a_{ij}=p_{jj}$，$i,j,r,s=1,\cdots,n$. 再取 $p_{11}=1$，$p_{jj}=a_{1j}$，$p_{ij}=0(i\neq j)$，$i,j=1,\cdots,n$，就合乎要求.

12. 只须考虑 A 的 Jordan 标准形. 答案是：A 的零特征根对应的初

等因式只有一次的，非零特征根或者是 $p-1$ 次单位根，或者是某个 p^a-1 次单位根，($1<a\leqslant t$，这里 t 是 A 的不同特征根的个数)当是p^a-1次单位根时，相应的特征根 $\lambda_{k1},\cdots,\lambda_{ka}$ 的同一方幂的初等因式有相同的个数，这里 $\lambda_{k1}^p=\lambda_{k2},\cdots,\lambda_{ka}^p=\lambda_{k1}$.

13. 先证明只须对 A，B都是 Jordan 标准形进行讨论. 再将X适当分块，把矩阵方程 $AX=XB$ 转化成一个矩阵方程组，从其中任取一个方程，通过仔细计算去研究. 答案是：有非零解X的充分必要条件是A，B有相同的特征根。

14. 取 $P=\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \omega^{(n-1)2} & \cdots & \omega^{(n-1)(n-1)} \end{pmatrix}$，其中 $\omega=\cos\frac{2\pi}{n}+\sqrt{-1}\sin\frac{2\pi}{n}$.

15. (i) 用归纳法. (ii) 参见第 14 题，但要用分块方阵.

习题 13.3

1. $X(t)=\exp(tA)$.

2. 用到第 1 题.

3. 利用矩阵幂级数.

4. 先证明不妨设 A 为 Jordan 块，然后直接计算.

5. 利用矩阵幂级数.

6. 存在酉方阵U，使 $O=U\mathrm{diag}(\lambda_1,\cdots,\lambda_n)\overline{U}'$，其中 $\lambda_j>0$，$j=1,\cdots,n$. 记 $S=\mathrm{diag}(\ln\lambda_1,\cdots,\ln\lambda_n)$，只须证明 $K=-\sqrt{-1}US\overline{U}'$是实斜对称方阵.

7. 前一问用定理 13.3.6. 后两问利用纯函数的定义化为不妨设 X 是 Jordan 块，再通过计算证明.

8. 利用第 7 题的结论.

9. 用第 8 题及定理 13.3.6. 可分 $l>0$ 和$l<0$ 两种情况讨论.

10. 化为证明存在 n 阶非异方阵 T，使 $TN^{(n)}T^{-1}=S$. 取

$$T=\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}-\sqrt{-1}\begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}$$

便可. 最后一问用到定理 13.3.8.

索　引

二　划

三　划

四　划

五 划

六 划

七 划

八　划

九 划

十 划

十一划

十二划

十三划

十四划

十五划

[G e n e r a l I n f o r m a t i o n]
书名= 线性代数与矩阵论
作者= 许以超
页数= 5 1 2
S S 号= 1 0 0 6 8 9 1 3
出版日期= 1 9 9 2 年0 5 月第1 版